

SNP genotyping and mapping in red raspberry

Development of a Flex-Seq SNP panel for raspberry (*Rubus idaeus* L.) and validation through linkage map construction and identification of QTL for several traits of agronomic importance to raspberry breeding

Davik J^{1*}, Zucchi P², Buti M³, Milne L⁴, Graham J⁵, Sargent DJ⁶

¹Department of Molecular Plant Biology, Norwegian Institute of Bioeconomy Research, Ås, Norway

² Sant'Orsola SCA, Department of Fruit Production, Via dell'Aeroporto 14, 38057, Pergine Valsugana (TN), Italy

³Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Florence, Italy

⁴Informational and Computational Sciences, The James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK

⁵Cell and Molecular Sciences, The James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK

⁶NIAB, 93 Lawrence Weaver Road, Cambridge CB3 0LE, UK.

*Author for correspondence

SNP genotyping and mapping in red raspberry

Abstract

High-throughput and reproducible genotyping platforms are critical for advancing genetic research and breeding in horticultural crops. Here, the development and validation of a custom single nucleotide polymorphism (SNP) panel using the Flex-Seq genotyping platform for red raspberry (*Rubus idaeus* L.) is described. SNPs were derived from existing linkage maps and RNA-seq data, resulting in a panel of 5,639 high-confidence, bi-allelic markers distributed across the seven chromosomes of the *R. idaeus* ‘Malling Jewel’ reference genome. The panel was used to genotype 457 red raspberry accessions including 161 individuals from a bi-parental mapping population (Paris×486), enabling the construction of high-density linkage maps and the identification of quantitative trait loci (QTL) for fruit size, leaf colour, plant vigour, and thorn density. Genome-wide association studies (GWAS) identified a major QTL for thornlessness on chromosome 4, co-locating with a candidate HOX3 gene, and multiple QTLs associated with anthocyanin biosynthesis genes for leaf colour. The SNP panel demonstrated utility for linkage mapping and trait association analyses, offering a powerful resource for marker-assisted selection and genetic improvement in red raspberry.

Key words: Single nucleotide polymorphism; Genome-wide association study; Quantitative trait locus; linkage mapping; *Rubus idaeus*; thornlessness; breeding; variety development; selection; genomics.

SNP genotyping and mapping in red raspberry

Introduction

Red raspberry (*Rubus idaeus* L.) is the most economically important species in the highly diverse Rosaceous genus *Rubus*. The species is prized primarily for its edible drupelets that develop into sweet, aromatic, usually red ‘berries’ which detach easily from their receptacles and are predominantly sold for fresh consumption throughout the world. The global production of red raspberries has steadily increased over the past 20 years, in part due to the development of high-quality varieties that can be produced throughout the temperate regions of the world, and produce high yields, delivering profitable returns for growers and retailers. To assist in the development of new red raspberry varieties, commercial breeding programmes are turning to the use of molecular markers to assist the selection process. Recently, several genome sequences for red raspberry have been published for the species [1-4], facilitating the development of large numbers of markers for use in high-throughput genotyping.

Rapid and reliable genotyping is one of the key foundations of any genetics study, and can be used to study diversity, develop genetic linkage maps [5, 6] and identify markers associated with traits of agronomic importance [7, 8]. High quality genotyping has been achieved in many crop species through the use of hybridisation-based genotyping microarrays, such as the Illumina Infinium arrays developed in apple [9], cherry [10], and peach [11], and the Affymetrix Axiom arrays developed in rose [12], apple [13] and strawberry [14], which produce reliable genotypes for many thousands, to hundreds of thousands of SNPs. Traditionally however, genotyping platform development has been the rate-limiting step in advancing the use of high-throughput SNP genotyping for understanding of the genetic control of agronomically-important traits in many crop species, due to prohibitive development costs.

Genotyping-by-sequencing (GBS) [15] is a relatively inexpensive method for the identification of large numbers of polymorphic SNP loci within a genome. It exploits the

SNP genotyping and mapping in red raspberry

distribution of restriction enzyme sites within a given genome to produce reduced-representation genomic libraries, and therefore requires no prior genomic information, using short-read sequencing to produce sequence data for SNP loci distributed genome-wide within any species. The technique has been used successfully in raspberry to develop SNP-based linkage maps of several mapping populations, either as the predominant marker type used, such as in the ‘Heritage’ × ‘Tulameen’ linkage map [16], or to ‘saturate’ linkage maps previously-produced using other marker types, such as in the case of the ‘Glen Moy’ × ‘Latham’ linkage map [17]. The ‘Glen Moy’ × ‘Latham’ linkage map has shown utility in several further studies permitting QTL for a range of traits to be identified, including fruit ripening [17], hyperspectral signatures correlated to several physical traits [18], crumbly fruit [19] and polyphenol metabolites [20]. However, there are many challenges with SNP genotyping using the GBS technique, including high levels of missing, or error-prone data, and most importantly, low reproducibility of genotyping between progenies.

Recently, several SNP genotyping platforms employing sequence-capture and high-throughput sequencing have been developed that are significantly more reproducible between experiments than GBS, but far more cost-effective and flexible than hybridisation-based genotyping arrays, and as such, are attractive for genetic investigation in crop species. These include the Flex-Seq platform, that has been successfully used to develop high density genotyping panels for blueberry [21, 22], and the single primer enrichment technology (SPET) genotyping that has been used to develop genotyping panels for tomato and aubergine [23] and lettuce [24]. Using these platforms, the cost of whole-genome SNP genotyping has reduced to the point whereby it is now cost-effective to employ the approach for almost any genetics study. Flex-Seq [25] is a SNP genotyping platform that employs sequence capture using targeted probes followed by Illumina short-read sequencing. The platform is claimed to provide highly accurate, reproducible data of a similar quality and consistency to data from genotyping arrays

SNP genotyping and mapping in red raspberry

[25] and since the platform relies on hybridisation of oligonucleotide probes, new markers can be added to existing SNP panels far more readily than with array-based genotyping platforms. Flex-Seq has several advantages over GBS; it requires prior knowledge of genomic variation of interest, which means it is highly targeted and customisable, ensuring specific regions of interest, or overall genome coverage can be achieved with a given SNP panel. It is also more robust to variation in DNA quality and as such, gives a more reproducible and reliable output in terms of sequence depth and representation, as well as between experiments than can be achieved with GBS [25].

For these reasons, in this investigation, the LGC Flex-seq SNP genotyping platform was employed to develop a SNP genotyping panel for red raspberry (*R. idaeus*). Previously-published molecular markers that had been mapped to genetic maps of several red raspberry varieties [16, 17] were used to create a framework SNP set spanning the ‘Malling Jewel’ reference genome sequence [4] and these were supplemented with robust SNPs characterised from RNASeq data of developing fruit tissues previously reported by [3]. The SNP panel was used to characterise a red raspberry diversity panel and a bi-parental mapping population, and QTL for several traits were identified to demonstrate the efficacy of the panel for genetics studies in *R. idaeus*.

Materials and Methods

Plant material and phenotyping

A total of 457 red raspberry genotypes comprising a diversity panel that included 27 red raspberry varieties (S1 File), 269 breeding selections from the Sant’Orsola raspberry breeding programme, and 161 progeny of an F1 mapping population were studied in the field

SNP genotyping and mapping in red raspberry

experiments. The bi-parental F_1 mapping population was raised from a controlled cross between the Sant'Orsola raspberry selections 'Paris' (female) and 'SO486' (male) (denoted Paris×486) was phenotyped in the same field trial. All raspberry germplasm was grown and maintained in substrate in a common breeding plot following standard cultivation practices at the Sant'Orsola S.C.A breeding centre at Altopiano della Vigolana, Trentino, Italy.

The field experiment followed an augmented design with regular control varieties. Subjective scores were taken regularly through each of the four experimental years using ordinal scales (0 – 10) for four horticultural traits; plant vigour, leaf colour, fruit size, and thorn density, the highest scores indicating more vigorous, darker leaves, larger and higher thorn density, respectively. Where necessary, data were transformed using Box-Cox transformation [26]. Each plot consisted of six clonal plants of each genotype. Phenotyping was performed over a total of four years and were analysed using ASReml-R [27] with entries as fixed and additional factors (e.g., years, plant cycle, and interaction terms) as random factors in order to get best linear unbiased estimators (BLUEs). As input for the quantitative trait linkage analysis (QTL) and the genome wide association analysis (GWAA) BLUEs across all years were used.

Marker development.

Selection of SNPs

Data for SNP marker development was derived from two sources. Firstly, markers that had previously identified using the Genotyping by Sequencing protocol of [15] and mapped in the 'Heritage' × 'Tulameen' (H×T) mapping population [28] and the 'Glen Moy' × 'Latham' mapping population [17] were retrieved from those datasets. Sequence Tag data containing the mapped SNPs were aligned to the 'Malling Jewel' reference sequence [4] using Blastn [29] and the SNP positions within the reference sequence were determined.

SNP genotyping and mapping in red raspberry

SNPs were then identified from red raspberry RNASeq data derived from the raspberry cultivars ‘Anitra’, ‘Glen Ample’, ‘Varnes’ and ‘Veten’ previously published by [3]. The quality of the RNASeq libraries was assessed using FastQC v0.11.9 [30], and poor quality reads and TruSeq adapters sequences were filtered out with Trimmomatic 0.36 [31] (parameters: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:5 TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:100). Filtered reads of all the libraries were mapped to the ‘Malling Jewel’ genome sequence [4] using the STAR version 2.7.3a [32] aligner in a two-pass mode to improve splice junction detection. In the first pass, reads from each sample were mapped to a STAR-generated genome index incorporating known splice junctions from the provided GTF annotation file [32]. The detected splice junctions from the first-pass alignment of each sample were then used to build a new second-pass STAR genome index for that specific sample. The same sample was then realigned using this refined index to improve mapping accuracy at exon-exon junctions. Aligned reads were sorted and duplicates marked using GATK [33] MarkDuplicates v4.4.0.0. Given the spliced nature of RNASeq data, GATK SplitNCigarReads was applied to preprocess reads before variant calling. Single-nucleotide polymorphisms (SNPs) and small insertions/deletions (indels) were identified using GATK HaplotypeCaller in GVCF mode for each sample, followed by joint genotyping with GATK GenotypeGVCFs to produce a final multi-sample VCF file. The raw VCF file was then filtered using bcftools v1.21-25-g9d314180 [34] based on quality thresholds, including variant quality (QUAL > 30), mapping quality (MQ > 40), read depth (DP > 10), minor allele frequency (AF > 0.05), and BaseQRankSum (-2 to 2), to retain high-confidence variant calls for downstream analysis. The physical positions of the SNPs identified from each dataset were used to create a BED file that was submitted to the LGC Flex-Seq panel design pipeline (LGC Genomics GmbH, Germany) for probe and adapter design, synthesis, and validation.

SNP genotyping and mapping in red raspberry

SNP genotyping

Approximately 50 mg of leaf material was harvested from each of the 457 genotypes, preserved using the LGC BioArk leaf collection kit and sent to LGC for DNA extraction using the oKtopure DNA extraction platform (LGC Genomics GmbH, Germany) and sequence data were generated using the Flex-Seq SNP panel and sequencing pipeline using the NexSeq 500 Illumina sequencing platform at LGC Genomics (LGC Genomics GmbH, Germany). In order to capture data from all SNPs within the sequenced amplicons, the raw reads were trimmed using Trimmomatic 0.39 [31] using the following settings: LEADING:5 TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:100, and were aligned using BWA-mem version 0.7.17-r1188 [35] to an indexed version of the ‘Malling Jewel’ raspberry reference genome [4]. The SNPs identified from the sequence data were called with the HaplotypeCaller in GATK [36]. Filtering was done with bcftools [34] keeping only biallelic SNPs, not accepting missing values below 0.1, requiring read quality to be above 30, retaining read depths below 300 and above 10, accepting BaseQRankSum below 3 and above -3, and finally, a minor allele frequency above 0.05. Following SNP identification, marker density was plotted from the genomic locations of SNPs in the ‘Malling Jewel’ reference genome [4] using the R package rMVP [37] to show the distribution of SNPs throughout the raspberry genome.

Linkage mapping

The SNP genotypes called in the 161 segregating full-sib family were further scrutinised to identify out-crossing, contaminated samples and rogue genotypes. Using an in-house R-code, the marker data was interrogated and ‘impossible genotypes’ (genotypes that could not have been derived from the mating of the two parental genotypes, i.e. the presence of genotype BB in the cross AA×AB) were flagged for each progeny. To account for a limited

SNP genotyping and mapping in red raspberry

amount of erroneous SNP calling from the sequence data, a threshold of more than five impossible genotypes per progeny was set and any progeny exceeding this value were removed from further analysis. Any remaining impossible genotypes in the retained progeny were set to missing values before linkage analysis, QTL mapping and GWAS were performed.

Joinmap 4.1 (Kyasma, NL) was used to group markers into linkage groups using the groupings tree function and a LOD threshold of 8.0. The genetic groupings were compared to the physical position of the markers on the raspberry ‘Malling Jewel’ reference genome and markers whose genetic linkage group did not match with their physical chromosome were removed from further analysis. The remaining markers were ordered by physical position along each chromosome and marker data imputed according to the methodology of [16].

From all segregating markers, 25 markers heterozygous in the male parent and 25 heterozygous in the female parent that were approximately evenly distributed throughout each chromosome based on physical position were selected from each chromosome, along with a further 25 markers from each chromosome that were heterozygous in both parents. These 525 segregating markers were combined with the 20 most significant markers identified from a Kruskal-Wallis analysis for each of the four phenotypes investigated (65 markers in total as some significant markers were common between traits) to give a total of 590 markers. Regression mapping, implemented in JoinMap 4.1 (Kyasma, NL) using default parameters was performed using the 590 markers to calculate a linkage map for both the male and female parents of the Paris×486 mapping population. The resultant linkage maps were plotted using MapChart 2.0 [38] with lines connecting common markers heterozygous and segregating in both parental maps (i.e. markers with the genotype AB×AB). Colinearity between genetic and physical positions of the mapped markers was determined by producing scatter plots of the female and male marker distribution, plotted using ggplot2 [39].

SNP genotyping and mapping in red raspberry

215 ***QTL analysis of traits segregating in the Paris×486 mapping progeny***

216 The best linear unbiased estimates calculated (BLUEs) for the mean phenotypic values
 217 calculated as described above were used as input for QTL analysis in the Paris×486 mapping
 218 population. A Kruskal-Wallis analysis was performed for each of the phenotypic trait datasets
 219 and the 20 most significant markers for each trait were included in the linkage map created
 220 using regression mapping. QTL analyses were performed using interval mapping implemented
 221 in MAPQTL 6.0 (Kyazma, NL) with a step size of 1.0 cM, from which percentage phenotypic
 222 variance explained and associated LOD values were calculated. A permutation test to
 223 determine the genome-wide LOD threshold for each trait was performed using 20,000
 224 permutations. Significant QTL identified for each of the four traits were plotted on the
 225 Paris×486 parental linkage maps using MapChart 2.0 [38].

226

227 **Genome-wide association (GWAS) analyses**

228 The best linear unbiased estimates (BLUEs) for plant vigour, leaf colour, fruit size, and
 229 thorn density calculated for the red raspberry accessions from the diversity panel that were both
 230 genotyped and phenotyped. The BLUEs were used as input for the GWAS analyses. A GWAS
 231 was performed using all heterozygous markers in the diversity panel. The GWAS was
 232 conducted using the single-locus mixed linear model [MLM, 40] and the multiloci Bayesian-
 233 information and linkage-disequilibrium iteratively nested keyway model [BLINK, 41]
 234 implemented in the GAPIT R Software package, version 3 [42]. To account for population
 235 structure and relatedness among entries, five principal components and the VanRaden [43]
 236 relationships matrix were included in the analyses. Manhattan plots showing the most
 237 significant associations on all seven red raspberry chromosomes were plotted using ggplot2
 238 [39].

SNP genotyping and mapping in red raspberry

239

240 Results

241 Plant material phenotyping

242 Phenotyping was performed for four traits; plant vigour, leaf colour, fruit size, and thorn
243 density in order to determine the efficacy of the SNP panel developed for genetic investigation
244 in red raspberry. A total of 457 accessions, including 161 F1 progeny of the Paris×486 mapping
245 population, 27 cultivars, and 269 breeding lines were phenotyped. The distribution of
246 phenotypes scored approximated a normal distribution for plant vigour, leaf colour, and fruit
247 size. The distribution of thorn density, however, was highly skewed. Box-Cox transformation,
248 fitted the trait distribution for thorn density to a more normal distribution, and the transformed
249 data were used for subsequent association and QTL analyses (**Fig 1**). The histograms showing
250 the phenotypic distributions for the Paris×486 mapping population are shown in **S1 Fig**. The
251 best linear unbiased estimates (BLUEs) calculated for fruit size, leaf colour, plant vigour, and
252 thorn density in the red raspberry diversity panel and used for subsequent GWAS analysis are
253 given in **S3 File**.

254

255 **Fig 1.** The distributions of phenotypes in the diversity panel of 457 red raspberry accessions
256 studied in this investigation. Distributions of phenotypes for (A) plant vigour, (B) leaf colour,
257 (C) fruit size, and thorn density (D).

258

259

260

SNP genotyping and mapping in red raspberry

SNP panel design

A total of 12,983 SNPs were selected from the three sources described in the materials and methods and submitted to the LGC design pipeline. The SNPs were located to 7,104 1 kb windows and following removal of loci outside of the percentage GC range, probes were designed for 7,071 loci, which were predicted *in silico* to contain 8,974 SNPs.

SNP genotyping in red raspberry germplasm

DNA was extracted from leaf material of the 457 accessions studied in this investigation, and following hybridisation of the Flex-Seq probes, library construction and sequencing using the NexSeq 500 Illumina sequencing platform at LGC Genomics (LGC Genomics GmbH, Germany) a total of 816,693,446 mapped 150bp paired-end reads were retrieved from the 481 samples. The average number of mapped reads per sample was 1,697,907 (min = 98,992, max = 2,378,375) (**S2 File**). Following SNP calling a .vcf file was produced containing a total of 116,731 unfiltered SNPs. The number of SNPs retained at each stage of filtering is shown in **Table 1** and the corresponding bed file (**S4 File**). Following filtering, a total of 5,639 heterozygous bi-allelic SNP sites (**S5 File**) were identified derived from the 7,071 probes. Distribution of the 5,639 bi-allelic SNPs on the ‘Malling Jewel’ genome sequence is shown in **Fig 2**. The plot revealed eight large gaps of more than 5 Mb in length where no bi-allelic SNPs were identified. At least one large gap was identified on each of the seven chromosomes, and as such may correspond to the centromeric regions of the raspberry genome.

SNP genotyping and mapping in red raspberry

Table 1. The number of sequenced SNPs retained at each stage of the filtering process in *R. idaeus* selections.

Purpose	Call	Remaining markers
Filter variants with missing genotypes gt 0.1	bcftools filter -e 'QUAL < 30' gatk1.vcf -O v	116,731
Filter on read quality	bcftools filter -e 'QUAL < 30'	116,731
Filter on read depth	bcftools filter -i 'FORMAT/DP > 10'	116,731
Filter on BaseQRankSum*	bcftools filter -i 'BaseQRankSum < 0.3 && BaseQRankSum > -0.3'	28,472
Filter on allele frequency	bcftools view -i 'AF > 0.01'	8,323
Only biallelic snps	bcftools view -v snps --max-alleles 2	5,639
Only indels	bcftools view -v indels --max-alleles 2	785

Fig 2. A heatmap showing the distribution of 5,639 biallelic SNP markers in 1 Mbp windows across the seven chromosomes of the ‘Malling Jewel’ raspberry reference genome sequence.

Paris×486 Linkage map construction

A total of 4,881 SNPs were heterozygous between the parents of the Paris×486 mapping population, 1,098 were heterozygous only in the female parent (AB×AA), 2,084 were heterozygous only in the male parent (AA×AB), with the remaining 1,699 heterozygous in both parental genotypes (AB×AB). A total of 35 progeny returned ‘impossible’ genotypes following genotyping and were removed from further study, leaving a total of 126 progeny for linkage analysis. Following initial linkage analysis, removal of markers that did not cluster to the expected group based on their physical location on the ‘Malling Jewel’ genome and subsequent imputation, a total of 4,736 segregating markers were mapped to seven linkage groups representing the seven chromosomes of the red raspberry genome (**S2 Fig**). Regression mapping with the selected subset of 590 markers described in the materials and methods produced a male and female linkage map spanning the expected seven chromosome and

SNP genotyping and mapping in red raspberry

covering a genetic distance of 610.8 cM and 654.5 cM for the male and female maps respectively. Marker ordering was well-conserved between the male and female maps and collinearity to the ‘Malling Jewel’ reference genome was high (**Fig 3**).

Fig 3. A plot of the genetic vs. physical positions of the SNP markers mapped in the Paris×486 mapping population. Markers in blue are those mapped to the ‘Paris’ map whilst those in red are those mapped to the ‘486’ map.

QTL analysis of the four phenotypic traits in the Paris×486 mapping population

Following QTL analysis and permutation testing, a total of five significant QTL surpassing the LOD threshold of 5.0 were discovered for the four phenotypic traits investigated. One significant QTL for fruit size was identified on the distal end of Ri2, with the most significant marker explaining 17.9% of the phenotypic variation observed, one QTL was identified for plant vigour on the proximal end of Ri3 with the most significant marker explaining 23.7% of the phenotypic variation observed, two QTL were observed for leaf colour on the proximal end of Ri3A and the distal end of Ri5A, with the most significant markers explaining 22.6% and 24.8% of the phenotypic variation observed respectively, and a final QTL for thorn density explaining 32.1% of the phenotypic variation observed was identified on the proximal end of Ri3 (**Fig 4**). All QTL were identified in both the male and female maps and the most significant marker trait associations along with their respective percentage phenotypic variance explained and associated LOD values and the marker positions on the Paris×486 linkage map are given in **Table 2**.

SNP genotyping and mapping in red raspberry

Fig 4. Genetic linkage map and QTL summary for fruit size, leaf colour, plant vigour and thorn density derived from the raspberry Paris×486 mapping population. QTL are represented by bars (1-LOD interval) and extended lines (2-LOD interval). Colours served to distinguish between QTL for different traits: red – fruit size; blue – leaf colour; green – plant vigour; black – thorn density. A. represents the linkage map of the female ‘Paris’ parent, whilst B. represents the male ‘S0486’ parent. Linkage groups are to scale and represented in cM according to the scale bars to the left of each linkage map.

Table 2. The significant QTLs identified in the Paris×486 mapping population, along with the most closely associated SNP markers, their physical position on the ‘Malling Jewel’ genome sequence, their corresponding genetic position and the LOD and percentage of observed variance explained by each QTL.

Trait	Marker	Physical position (bp)	Linkage group	Genetic position (cM)	LOD	Variance explained (%)
Fruit size	c2_31392842	31,392,842	Ri2A/B	56.199	5.31	17.7
Fruit size	c2_31520020	31,520,020	Ri2A/B	57.051	5.39	17.9
Fruit size	c2_31751824	31,751,824	Ri2A/B	60.884	4.44	15
Leaf colour	c3_867579	867,579	Ri3A	0.966	6.52	21.2
Leaf colour	c3_1652814	1,652,814	Ri3A	1.309	6.81	22
Leaf colour	b3_1691853	1,691,853	Ri3B	4.059	6.97	22.6
Leaf colour	c3_1942363	1,942,363	Ri3A	3.074	6.52	21.2
Leaf colour	c5_13556169	13,556,169	Ri5A/B	83.58	7.79	24.8
Leaf colour	c5_13573690	13,573,690	Ri5A/B	83.596	7.79	24.8
Leaf colour	a5_14348691	14,348,691	Ri5A	85.633	7.95	25.4
Leaf colour	c5_14454153	14,454,153	Ri5A/B	85.754	7.72	24.7
Vigour	b3_2313205	2,313,205	Ri3B	8.757	6.83	22.1
Vigour	a3_2556608	2,556,608	Ri3A	7.204	6.85	22.2

SNP genotyping and mapping in red raspberry

Vigour	c3_2694951	2,694,951	Ri3A/B	7.212	6.85	22.2
Vigour	c3_2822294	2,822,294	Ri3A/B	8.021	7.4	23.7
Vigour	a3_2909967	2,909,967	Ri3A	7.17	6.86	22.2
Vigour	c3_3090358	3,090,358	Ri3A/B	8.402	6.8	22
Thorn density	a3_11758215	11,758,215	Ri3A	47.102	9.39	32.1
Thorn density	a3_12805561	12,805,561	Ri3A	49.564	9.51	31.9
Thorn density	b3_14248398	14,248,398	Ri3B	61.322	9.57	29.9
Thorn density	b3_14248464	14,248,464	Ri3B	61.289	9.63	30.1
Thorn density	b3_14306747	14,306,747	Ri3B	62.258	9.57	29.6
Thorn density	c3_15046391	15,046,391	Ri3A/B	54.038	9.12	28.4

GWAS for phenotypic traits in the red raspberry diversity panel

Significant associations were identified for three of the four phenotypic traits analysed using GAPIT: fruit size, leaf colour, and thorn density (**Fig 5**). A highly significant association was observed for thorn density on chromosome 4, with peaks of lower significance on chromosome 5 and 7, a significant association was observed for fruit size on chromosome 1, three significant associations were observed on chromosomes 3, 4 and 5 for leaf colour, whilst no significant associations were identified for plant vigour. The most significant association observed was for thorn density, with a signal peak on chromosome 4 at 33,160,893 bp explaining 34.8% of the phenotypic variation observed.

Fig 5. Manhattan plots of markers significantly associated with fruit size, leaf colour, and thorn density in the red raspberry diversity panel used in this investigation, constructed against the ‘Malling Jewel’ reference genome sequence showing marker-trait associations exceeding the $-\log_{10}(p)$ genome-wide threshold of ~ 7.5 (red dashed line). The suggestive threshold of $-\log_{10}(p) \sim 5$ is also indicated with the blue dotted line.

SNP genotyping and mapping in red raspberry

Candidate gene identification

The genomic regions surrounding each of the QTL were interrogated and the genes they contained were retrieved from the data of [4] and several of the QTL identified were associated with candidate genes. The most significant GWAS association with thorn density was at 33,160, 893 bp in the ‘Malling Jewel’ reference genome sequence. This marker was found to be in close proximity to the predicted gene jg27953.t1, at position 33,102,286 bp to 33,104,149 bp in the ‘Malling Jewel’ reference genome sequence that encoded a predicted homeobox-leucine zipper (HOX3) protein (**S6 File**). The most significant QTL for leaf colour were identified on LG3 and LG5 in both the mapping population and the GWAS study, with the most significant markers located at 1,652,814 bp and 2,822,294 bp in the mapping study and GWAS respectively on LG3 and at 9,150,237 bp and 11,495,066 bp in the mapping study (male linkage map) and GWAS respectively. Candidate genes were identified in all four regions associated with anthocyanin biosynthesis. On chromosome 3, predicted gene jg346.t1 was identified, located at 1,450,849 bp to 1,452,587 bp which encodes for a predicted anthocyanidin reductase gene, whilst four gene predictions jg637.t1, jg638.t1, jg639.t1, and jg640.t1 were located between 2,845,810 bp and 2,854,683 bp on the ‘Malling Jewel’ reference genome sequence that all encode for predicted anthocyanidin 3-O-glucoside 2"-O-glucosyltransferase-like genes. On chromosome 5, predicted gene jg19747.t1 was located at 7,905,815 bp to 7,907,502 bp in the ‘Malling Jewel’ reference genome sequence that encodes for anthocyanidin synthase, whilst predicted gene jg20417.t1 was located between 11,965,651 bp and 11,967,018 bp that is predicted to encode an anthocyanidin 3-O-glucosyltransferase 7-like gene.

Discussion

The main objective of this study was the development of a SNP marker genotyping panel for *R. idaeus* of broad utility to genetics studies in the species. To this end, a total of

SNP genotyping and mapping in red raspberry

5,639 SNP markers heterozygous in a diversity panel of 457 red raspberry genotypes was developed and validated through GWAS, and the development of a genetic linkage map for a bi-parental mapping population. Marker density was similar to other SNP panels that have been developed with analogous technologies for other plant species such as blueberry [21] and comparable to other SNP-based linkage mapping approached in red raspberry such as Genotyping by Sequencing [16] [17], but with the advantage of being proven to be highly reproducible when applied to highly heterozygous plant genomes [22]. Marker distribution throughout the *R. idaeus* genome ensured that, whilst some regions contained a higher abundance of heterozygous SNPs, the majority of the coding portion of the genome was represented, whilst marker density was low within the repetitive regions of the genome [4]. However, due to the adaptable nature of the FlexSeq genotyping platform, additional SNPs from the highly repetitive regions of the red raspberry genome could be added for future genotyping studies using the marker set presented here.

Using phenotypic data collected over several seasons, association studies revealed significant QTL for several traits in both a diversity panel of red raspberry genotypes and a bi-parental mapping population. The most significant QTL identified was for thorn density, a trait first mapped in red raspberry by Molina-Bravo et al [44], and in tetraploid blackberry, a close relative of red raspberry by Castro et al. [45], both of whom mapped the trait to linkage group 4 of the respective genomes. In the diversity panel, a single major QTL was identified on chromosome 4 associated with a SNP physically located at 33,160, 893 bp in the ‘Malling Jewel’ reference genome sequence. Genetic studies performed recently in tetraploid blackberry, identified a genetic region controlling the ‘prickle-free’ trait in the blackberry cultivar ‘Merton Thornless’ [46]. This locus was associated with a SNP marker located in an LD block at 33.64 Mbp on the blackberry genome sequence [47] and the authors revealed a potential causal mutation in a homologue of the HOX3 gene located between 33,599,420 and

SNP genotyping and mapping in red raspberry

33,601,244 bp in the blackberry genome, 35 kb from the peak marker. Scrutiny of the predictions in the region surrounding the peak SNP for thorn density in this investigation revealed a candidate gene *cg27953.t1* located between 33,102,286 bp and 33,104,149 bp on the ‘Malling Jewel’ reference genome sequence, 4,607 bp from the peak SNP, that encoded a predicted homeobox-leucine zipper (HOX3) gene, suggesting the QTL in red raspberry and blackberry may be controlled by the same gene homologue. Interestingly, the bi-parental mapping population did not reveal any thornless offspring and segregated only for thorn density. The QTL revealed in this population was located on chromosome 3, between 11-15 Mbp, and as such is controlled by a different gene to that controlling thornlessness on chromosome 4.

Candidate genes were also discovered in the QTL regions controlling leaf colour on chromosome 3 and chromosome 5 in both the diversity panel and the mapping population. On chromosome 3 in the biparental mapping population predicted gene *cg346.t1*, located between 1,450,849 bp and 1,452,587 bp, 200,227 bp from the peak SNP is predicted to encode for anthocyanidin reductase. Recently, loss of function mutations in anthocyanin reductase, a key gene in the proanthocyanidin synthesis pathway, have been shown to activate anthocyanin synthesis in strawberry, a close relative of red raspberry, with EMS induced mutants displaying increased anthocyanin accumulation during fruit development and higher anthocyanin content in ripe fruits compared to wild-type clones [48]. On chromosome 3 in the diversity panel four genes *cg637.t1*, *cg638.t1*, *cg639.t1*, and *cg640.t1* were located between 2,845,810 bp and 2,854,683 bp, 23,516 bp from the peak SNP, all encoding a predicted anthocyanidin 3-O-glucosyltransferase-like gene. The peak SNP for the significant QTL for leaf colour on chromosome 5 in the diversity panel was located at 11,495,066 bp, 470,585 bp from the predicted gene *cg20417.t1* (located between 11,965,651 bp and 11,967,018 bp) encoding an anthocyanidin 3-O-glucosyltransferase 7-like gene. UFGT enzymes are critical for anthocyanin

SNP genotyping and mapping in red raspberry

synthesis, acylation, and glucosylation in horticultural plants, as well as in the stabilization of anthocyanin pigments [49], and anthocyanidin-3-O-glucosyltransferases (UFGTs) have been shown to play a key role in the biosynthesis of stable anthocyanins, with anthocyanin accumulation significantly reduced in UFGTs mutants in *Arabidopsis* [50]. The peak SNP for the significant QTL for leaf colour on chromosome 5 was located at 9,150,237 bp in the mapping study, whilst gene jg19747.t1, encoding a predicted anthocyanidin synthase was located between 7,905,815 bp and 7,907,502 bp, 1,244,422 bp from the peak SNP in the ‘Malling Jewel’ reference genome sequence. Recently, mutations in anthocyanidin synthase were implicated in the loss of fruit colour in the red raspberry mutants ‘Anne’ [51] and ‘Varnes’ [3].

Whilst no causal mutations were identified in this study through linkage mapping or GWAS, the results demonstrate the efficacy of the SNPs in identifying QTL for agronomic traits and this the potential utility of the SNP panel developed for studies into the genetic basis of other important traits in red raspberry. It also demonstrates the potential for the SNP panel for use in the development of tools for marker assisted breeding, and for genomic prediction and selection, as has been shown in other species such as strawberry [52, 53] and chickpea (*Cicer arietinum* L.) [54], especially for complex polygenic traits such as yield and disease resistance which are of significant importance to commercial raspberry cultivation.

Conclusions

The SNP panel presented here has been shown to effective at developing linkage maps, and detecting QTL for traits of agronomic importance in red raspberry in both a diversity panel and a bi-parental mapping population. As such, the panel could assist in significantly accelerating the development of new raspberry cultivars with improved traits, such as enhanced disease resistance, better fruit quality, and increased yield, through the application of SNP

SNP genotyping and mapping in red raspberry

markers to breeding and selection. It will also be of utility for characterising and conserving genetic diversity, as well as for safeguarding the genetic security of commercial raspberry varieties through genetic fingerprinting, and as such should prove to be a valuable resource for researchers and breeders in the global raspberry community.

Supplementary Files

S1 Fig. The distributions of phenotypes in the Paris×486 mapping population. Distributions of phenotypes for (A) plant vigour, (B) leaf colour, (C) fruit size, and (D) thorn density.

S1 File. The 27 red raspberry varieties included in the diversity panel in this investigation.

S2 File. Summary of sequencing reads, including the number of adapter-clipped reads received from the sequencing facility (LGC), the number of reads retained after Trimmomatic trimming, and the number of reads successfully mapped to the Malling Jewel genome.

S2 Fig. Maximum Likelihood linkage map for the 4,736 segregating markers that located to the expected chromosome according to physical position in the Paris×486 mapping population. Markers segregating AB×AB that were mapped to both male and female maps are linked by a solid line between linkage maps.

S3 File. Best linear unbiased estimates (BLUEs) calculated for fruit size, leaf colour, plant vigour, and thorn density from the red raspberry diversity panel used for GWAS in this investigation.

S4 File. The BED file corresponding to the positions of the 5,369 SNP markers used in the current investigation.

S5 File. The numeric genotypes used in the current investigation. A ‘0’ indicates homozygous reference, a ‘1’ indicates a heterozygous, and a ‘2’ a homozygous alternate.

SNP genotyping and mapping in red raspberry

477 **S6 File.** Candidate genes retrieved from the ‘Malling Jewel’ reference genome sequence within
478 QTL spanning intervals from the GWAS and regression mapping analyses performed in this
479 study.

SNP genotyping and mapping in red raspberry

480 References

- 481 1. Davik J, Røen D, Lysøe E, Buti M, Rossman S, Alsheikh M, et al. A chromosome-
482 level genome sequence assembly of the red raspberry (*Rubus idaeus* L.). Plos one.
483 2022;17(3):e0265096.
- 484 2. Zhang H, Li W, Li G, Liu J, Chen H, Zhang C, et al. The High-Quality Genome
485 Sequencing and Analysis of Red Raspberry (*Rubus idaeus* L.). International Journal of
486 Genomics. 2024;2024(1):9271183. doi: <https://doi.org/10.1155/2024/9271183>.
- 487 3. Sargent DJ, Buti M, Martens S, Pugliesi C, Aaby K, Røen D, et al. A CACTA-like
488 transposon in the Anthocyanidin synthase 1 (Ans-1) gene is responsible for apricot fruit
489 colour in the raspberry (*Rubus idaeus*) cultivar 'Varnes'. PloS one. 2025;20(2):e0318692.
- 490 4. Price RJ, Davik J, Fernández Fernández F, Bates HJ, Lynn S, Nellist CF, et al.
491 Chromosome-scale genome sequence assemblies of the 'Autumn Bliss' and 'Malling
492 Jewel' cultivars of the highly heterozygous red raspberry (*Rubus idaeus* L.) derived from
493 long-read Oxford Nanopore sequence data. Plos one. 2023;18(5):e0285756.
- 494 5. Sargent DJ, Yang Y, Surbanovski N, Bianco L, Buti M, Velasco R, et al. HaploSNP
495 affinities and linkage map positions illuminate subgenome composition in the octoploid,
496 cultivated strawberry (*Fragaria x ananassa*). Plant Sci. 2016;242:140-50.
- 497 6. Di Pierro EA, Gianfranceschi L, Di Guardo M, Koehorst-van Putten HJJ, Kruisselbrink
498 JW, Longhi S, et al. A high-density, multi-parental SNP genetic map on apple validates a
499 new mapping approach for outcrossing species. Horticulture Research. 2016;3. doi:
500 10.1038/hortres.2016.57.
- 501 7. Li J, Eltaher S, Freeman B, Singh S, Ali GS. Comprehensive genetic diversity and
502 genome-wide association studies revealed the genetic basis of avocado fruit quality traits.
503 Frontiers in Plant Science. 2024;15. doi: 10.3389/fpls.2024.1433436.
- 504 8. Font i Forcada C, Guajardo V, Chin-Wo SR, Moreno MÁ. Association Mapping
505 Analysis for Fruit Quality Traits in *Prunus persica* Using SNP Markers. Frontiers in Plant
506 Science. 2019;9. doi: 10.3389/fpls.2018.02005.
- 507 9. Bianco L, Cestaro A, Sargent DJ, Banchi E, Derdak S, Di Guardo M, et al.
508 Development and Validation of a 20K Single Nucleotide Polymorphism (SNP) Whole
509 Genome Genotyping Array for Apple (*Malus × domestica* Borkh). PLOS ONE.
510 2014;9(10):e110377. doi: 10.1371/journal.pone.0110377.
- 511 10. Vanderzande S, Zheng P, Cai L, Barac G, Gasic K, Main D, et al. The cherry 6+9K
512 SNP array: a cost-effective improvement to the cherry 6K SNP array for genetic studies.
513 Scientific Reports. 2020;10(1):7613. doi: 10.1038/s41598-020-64438-x.
- 514 11. Verde I, Bassil N, Scalabrin S, Gilmore B, Lawley CT, Gasic K, et al. Development
515 and Evaluation of a 9K SNP Array for Peach by Internationally Coordinated SNP Detection
516 and Validation in Breeding Germplasm. PLOS ONE. 2012;7(4):e35668. doi:
517 10.1371/journal.pone.0035668.
- 518 12. Koning-Boucoiran CFS, Esselink GD, Vukosavljev M, van 't Westende WPC, Gitonga
519 VW, Krens FA, et al. Using RNA-Seq to assemble a rose transcriptome with more than
520 13,000 full-length expressed genes and to develop the WagRhSNP 68k Axiom SNP array for
521 rose (*Rosa* L.). Frontiers in Plant Science. 2015;6. doi: 10.3389/fpls.2015.00249.
- 522 13. Bianco L, Cestaro A, Linsmith G, Muranty H, Denancé C, Théron A, et al.
523 Development and validation of the Axiom®Apple480K SNP genotyping array. The Plant
524 Journal. 2016;86(1):62-74. doi: <https://doi.org/10.1111/tbj.13145>.
- 525 14. Hardigan MA, Feldmann MJ, Lorant A, Bird KA, Famula R, Acharya C, et al. Genome
526 synteny has been conserved among the octoploid progenitors of cultivated strawberry over
527 millions of years of evolution. Frontiers in Plant Science. 2020;10(1789). doi:
528 10.3389/fpls.2019.01789.
- 529 15. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust,
530 simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE.
531 2011;6:e19379.

SNP genotyping and mapping in red raspberry

16. Ward JA, Bhangoo J, Fernandez-Fernandez F, Moore P, Swanson JD, Viola R, et al. Saturated linkage map construction in *Rubus idaeus* using genotyping by sequencing and genome-independent imputation. BMC Genomics. 2013;14.
17. Hackett CA, Milne L, Smith K, Hedley P, Morris J, Simpson CG, et al. Enhancement of Glen Moy x Latham raspberry linkage map using GbS to further understand control of developmental processes leading to fruit ripening. BMC Genet. 2018;19(1):59. doi: 10.1186/s12863-018-0666-z.
18. Williams D, Hackett CA, Karley A, McCallum S, Smith K, Britten A, et al. Seeing the wood for the trees: hyperspectral imaging for high throughput QTL detection in raspberry, a perennial crop species. Fruit Research. 2021;1(1):1-11. doi: 10.48130/FruRes-2021-0007.
19. Scolari LM, Hancock RD, Hedley PE, Morris J, Smith K, Graham J. Combining QTL Mapping and Gene Expression Analysis to Elucidate the Genetic Control of 'Crumbly' Fruit in Red Raspberry (*Rubus idaeus* L.). Agronomy. 2021;11(4):794. PubMed PMID: doi:10.3390/agronomy11040794.
20. McDougall GJ, Allwood JW, Dobson G, Austin C, Verrall S, Alexander CJ, et al. Quantitative trait loci mapping of polyphenol metabolites from a 'Latham' x 'Glen Moy' red raspberry (*Rubus idaeus* L) cross. Metabolomics. 2023;19(8):71. doi: 10.1007/s11306-023-02033-7.
21. Qi X, Ogden EL, Bostan H, Sargent DJ, Ward J, Gilbert J, et al. High-Density Linkage Map Construction and QTL Identification in a Diploid Blueberry Mapping Population. Frontiers in Plant Science. 2021;12. doi: 10.3389/fpls.2021.692628.
22. Clare SJ, Driskill M, Millar TR, Chagné D, Montanari S, Thomson S, et al. Development of a targeted genotyping platform for reproducible results within tetraploid and hexaploid blueberry. Frontiers in Horticulture. 2024;2. doi: 10.3389/fhort.2023.1339310.
23. Barchi L, Acquadro A, Alonso D, Aprea G, Bassolino L, Demurtas O, et al. Single Primer Enrichment Technology (SPET) for High-Throughput Genotyping in Tomato and Eggplant Germplasm. Frontiers in Plant Science. 2019;10. doi: 10.3389/fpls.2019.01005.
24. Tripodi P, Beretta M, Peltier D, Kalfas I, Vasilikiotis C, Laidet A, et al. Development and application of Single Primer Enrichment Technology (SPET) SNP assay for population genomics analysis and candidate gene discovery in lettuce. Frontiers in Plant Science. 2023;14. doi: 10.3389/fpls.2023.1252777.
25. Jaimes H, Londoño A, Saavedra-Díaz C, Trujillo-Montenegro JH, López-Gerena J, Riascos JJ, et al. Sequencing vs. amplification for the estimation of allele dosages in sugarcane (*Saccharum* spp.). Applications in Plant Sciences. 2024;12(5):e11574.
26. Box GE, Cox DR. An analysis of transformations. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1964;26(2):211-43.
27. Butler DG, Cullis BR, Gilmour AR, Gogel BJ, Thompson R. ASReml-R Reference Manual Version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK. 2018.
28. Ward JA, Ponnala L, Weber CA. Strategies for Transcriptome Analysis in Nonmodel Plants. Am J Bot. 2012;99(2):267-76.
29. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):1-9.
30. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data 2010. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
31. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina Sequencing Data. Bioinformatics2014. p. 1-7.
32. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2012;29(1):15-21. doi: 10.1093/bioinformatics/bts635.
33. Van der Auwera GA, O'Connor BD. Genomics in the cloud: using Docker, GATK, and WDL in Terra: O'Reilly Media; 2020.
34. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2):giab008.
35. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:13033997v1 [q-bioGN]. 2013.

SNP genotyping and mapping in red raspberry

36. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, et al. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*. 2013;43(1):11.0.1-0.33. doi: <https://doi.org/10.1002/0471250953.bi1110s43>.
37. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, et al. rMVP: A Memory-Efficient, Visualization-Enhanced, and Parallel-Accelerated Tool for Genome-Wide Association Study. *Genomics, Proteomics & Bioinformatics*. 2021;19(4):619-28. doi: 10.1016/j.gpb.2020.10.007.
38. Voorrips RE. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J Hered*. 2002;93(1):77-8. doi: 10.1093/jhered/93.1.77.
39. Wickham H. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York, 2016. 2016.
40. Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet*. 2006;38:203-8.
41. Huang M, Liu X, Zhou Y, Summers RM, Zhang Z. BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience*. 2019;8(2):giy154.
42. Wang J, Zhang Z. GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genomics, Proteomics and Bioinformatics*. 2021;19(4):629-40.
43. VanRaden P. Efficient methods to compute genomic predictions. *J Dairy Sci*2008. p. 4414-23.
44. Molina-Bravo R, Fernandez GE, Sosinski BR. Quantitative trait locus analysis of tolerance to temperature fluctuations in winter, fruit characteristics, flower color, and prickle-free canes in raspberry. *Mol Breed*. 2014;33(2):267-80. doi: 10.1007/s11032-013-9947-4.
45. Castro P, Stafne ET, Clark JR, Lewers KS. Genetic map of the primocane-fruiting and thornless traits of tetraploid blackberry. *Theor Appl Genet*. 2013;126(10):2521-32. doi: 10.1007/s00122-013-2152-3.
46. Johns CA, Silva A, Chizk TM, Nelson L, Clark JR, Aryal R, et al. Genetic Control of Prickles in Tetraploid Blackberry. *G3 Genes|Genomes|Genetics*. 2025. doi: 10.1093/g3journal/jkaf065.
47. Brūna T, Aryal R, Dudchenko O, Sargent DJ, Mead D, Buti M, et al. A chromosome-length genome assembly and annotation of blackberry (*Rubus argutus*, cv. "Hillquist"). *G3 Genes|Genomes|Genetics*. 2022;13(2). doi: 10.1093/g3journal/jkac289.
48. Xu P, Li M, Ma C, Li X, Bai P, Lin A, et al. Loss-of-function mutation in anthocyanidin reductase activates the anthocyanin synthesis pathway in strawberry. *Molecular Horticulture*. 2024;4(1):33. doi: 10.1186/s43897-024-00106-2.
49. Muhammad N, Luo Z, Yang M, Li X, Liu Z, Liu M. The joint role of the late anthocyanin biosynthetic UFGT-encoding genes in the flowers and fruits coloration of horticultural plants. *Scientia Horticulturae*. 2022;301:111110. doi: <https://doi.org/10.1016/j.scienta.2022.111110>.
50. Yonekura-Sakakibara K, Fukushima A, Nakabayashi R, Hanada K, Matsuda F, Sugawara S, et al. Two glycosyltransferases involved in anthocyanin modification delineated by transcriptome independent component analysis in *Arabidopsis thaliana*. *The Plant Journal*. 2012;69(1):154-67. doi: <https://doi.org/10.1111/j.1365-313X.2011.04779.x>.
51. Rafique MZ, Carvalho E, Stracke R, Palmieri L, Herrera L, Feller A, et al. Nonsense mutation inside anthocyanidin synthase gene controls pigmentation in yellow raspberry (*Rubus idaeus* L.). *Frontiers in plant science*. 2016;7:238929.
52. Werner CR, Gaynor RC, Sargent DJ, Lillo A, Gorjanc G, Hickey JM. Genomic selection strategies for clonally propagated crops. *Theor Appl Genet*. 2023;136(4):74.
53. Sleper J, Tapia R, Lee S, Whitaker V. Within-family genomic selection in strawberry: Optimization of marker density, trial design, and training set composition. *The Plant Genome*. 2025;18(1):e20550. doi: <https://doi.org/10.1002/tpg2.20550>.
54. Lin Z, Li Y, Riaz A, Sudheesh S, Yazdifar S, Atieno J, et al. Assessing the utility of genomic selection to breed for durable *Ascochyta* blight resistance in chickpea. *The Plant Genome*. 2025;18(2):e70023. doi: <https://doi.org/10.1002/tpg2.70023>.

SNP genotyping and mapping in red raspberry

642

643

644

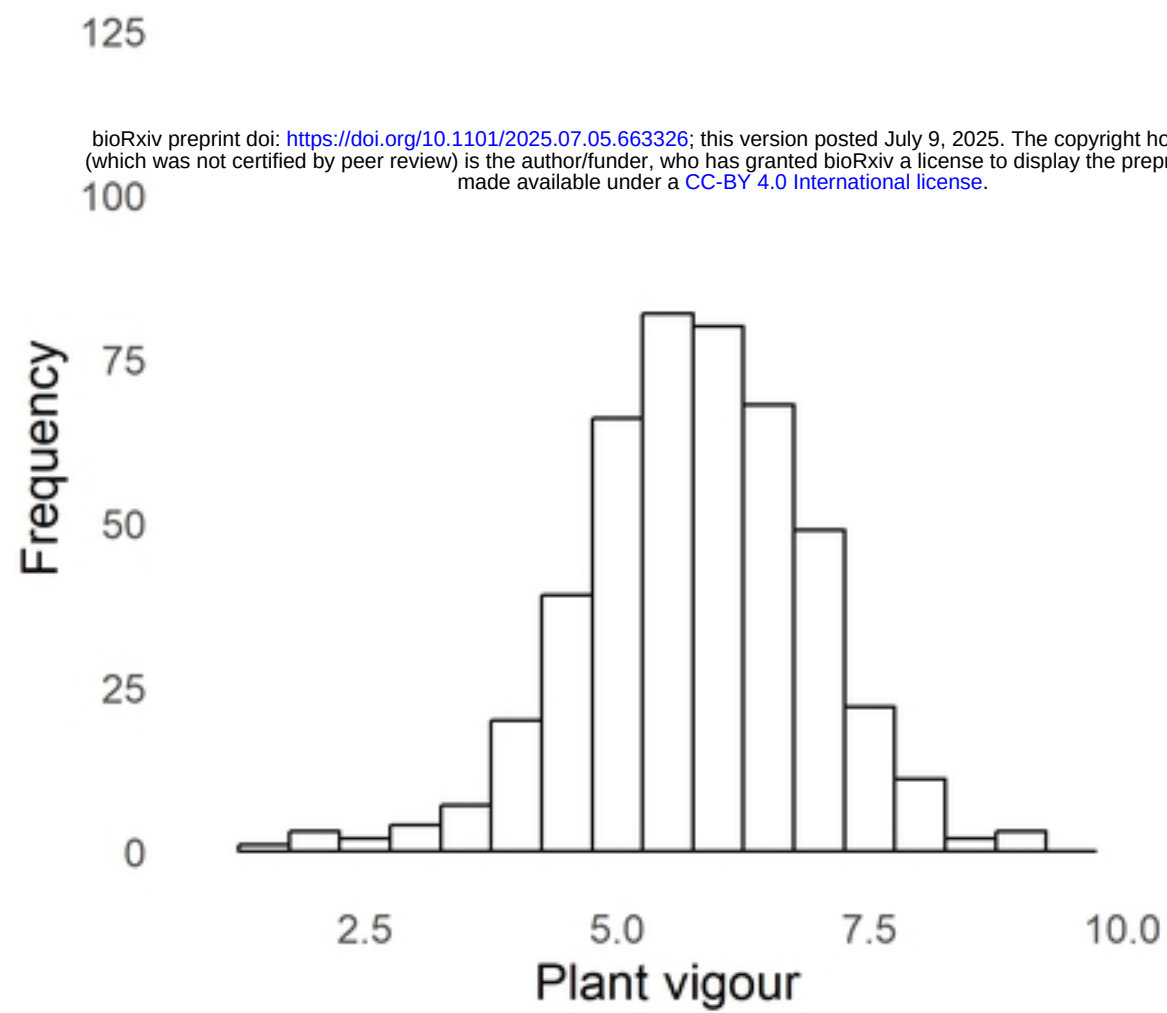
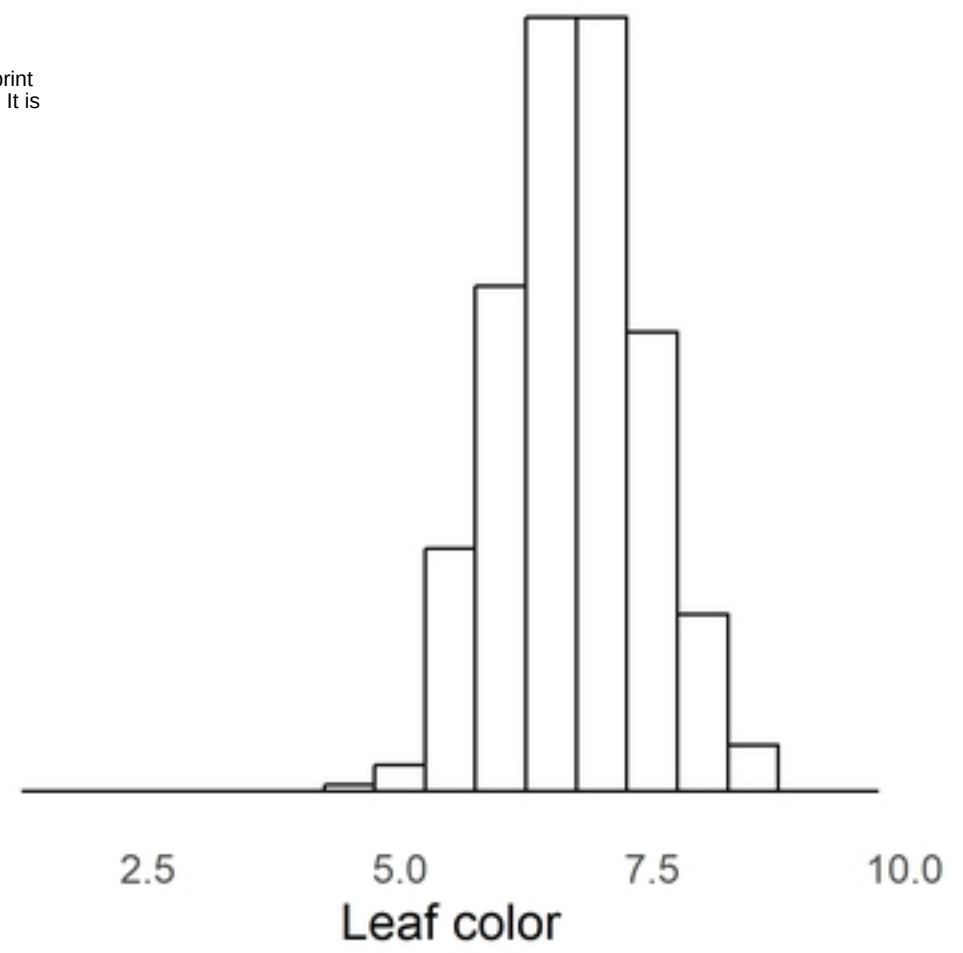
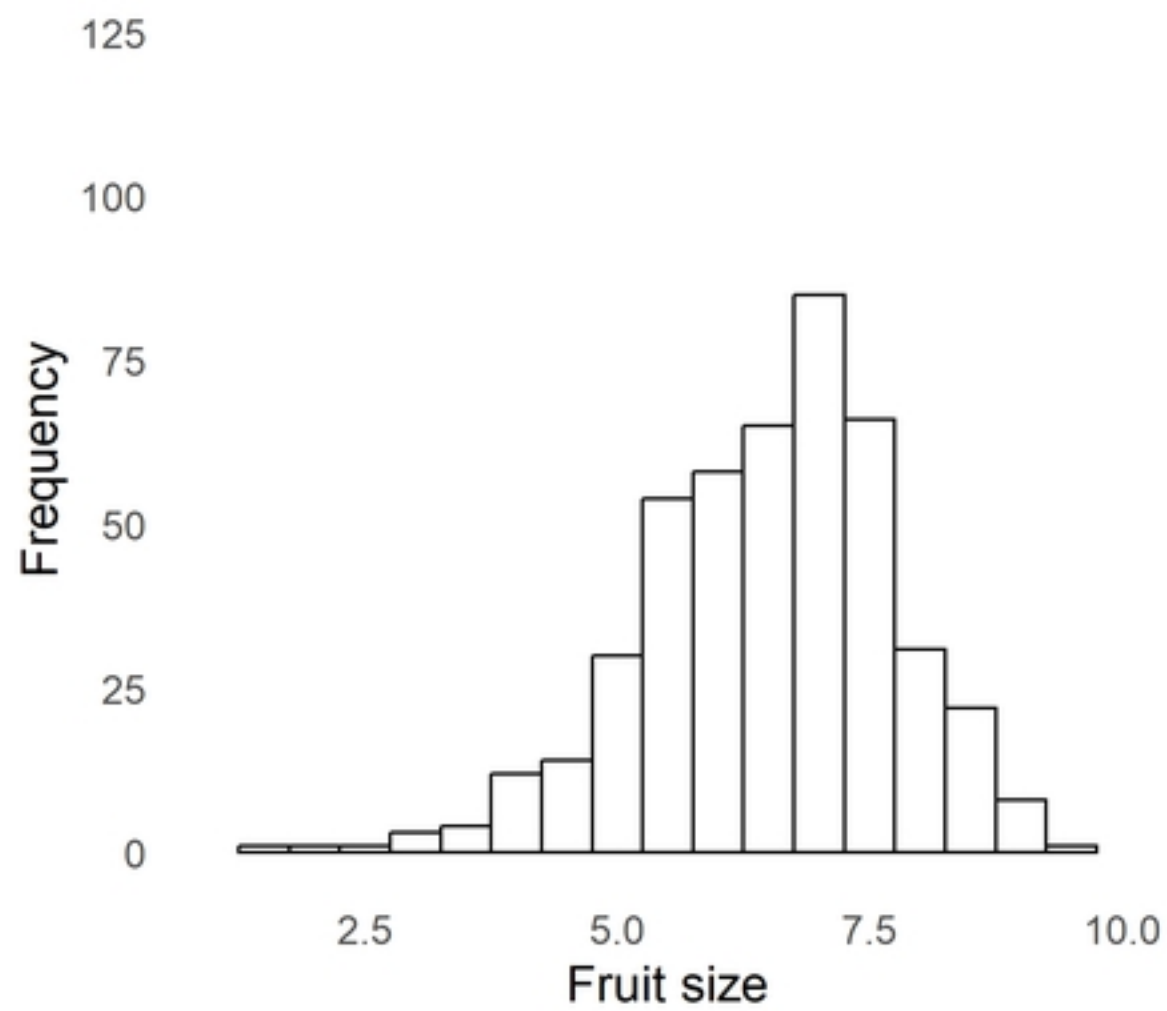
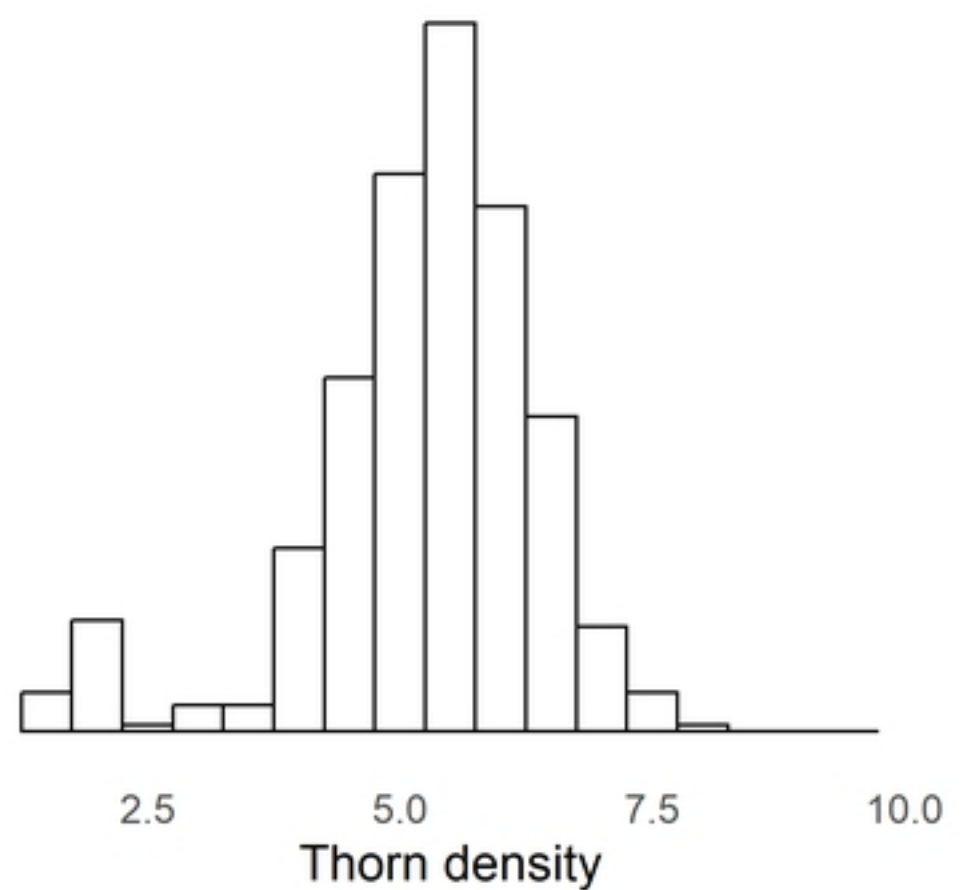
A**B****C****D**

Figure1 Phenotype distributions

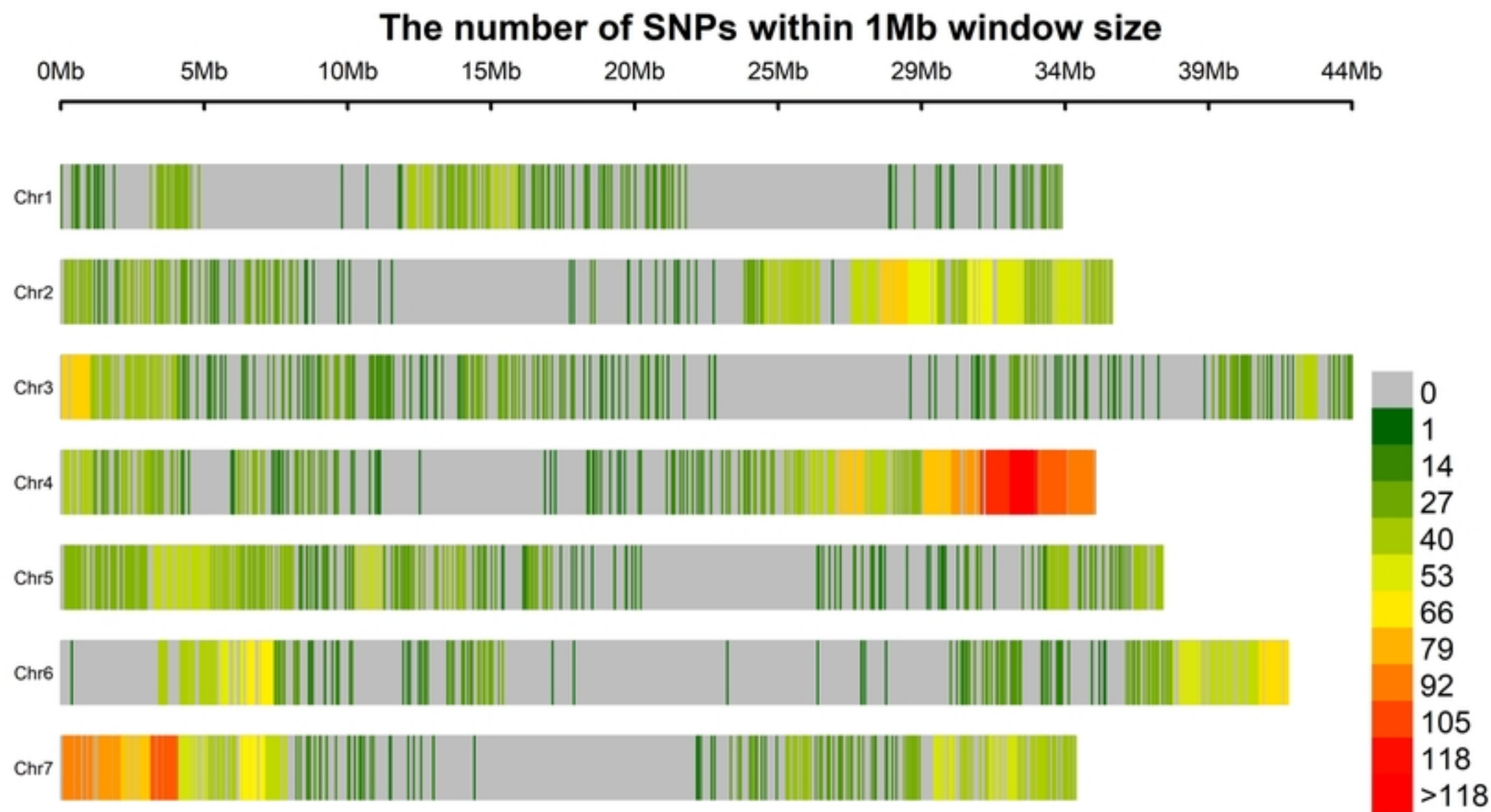
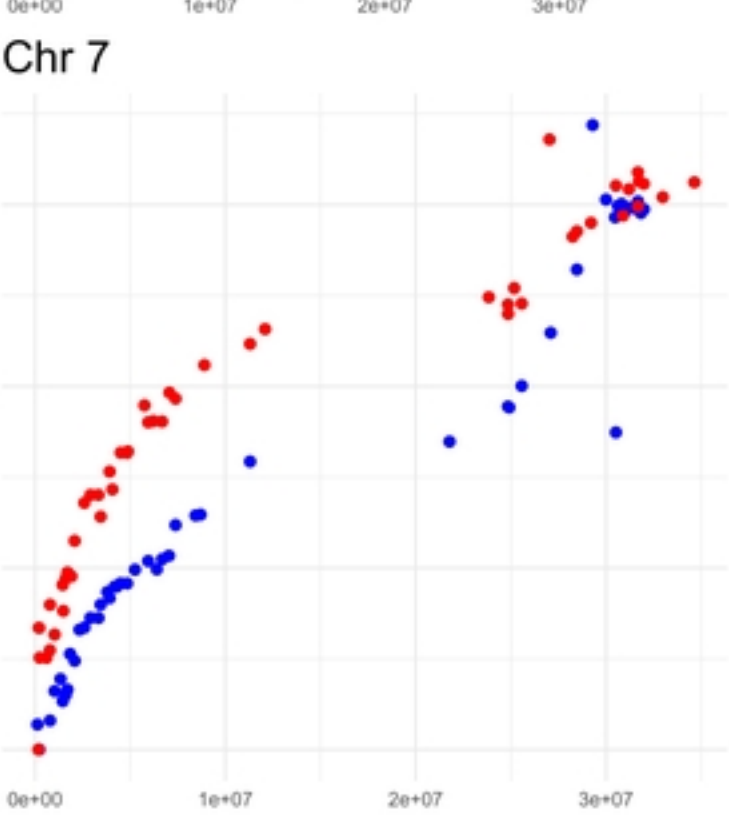
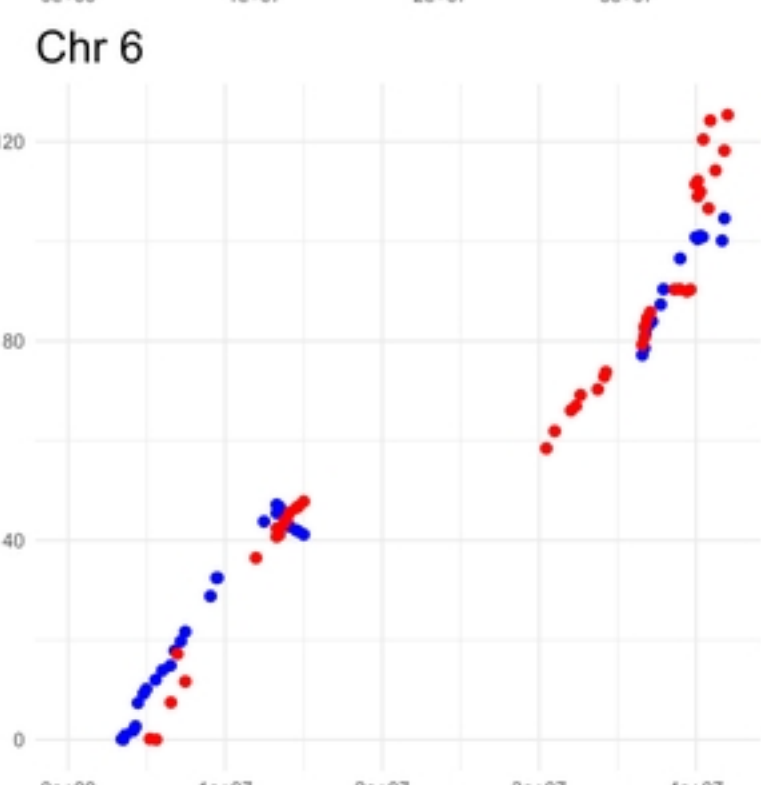
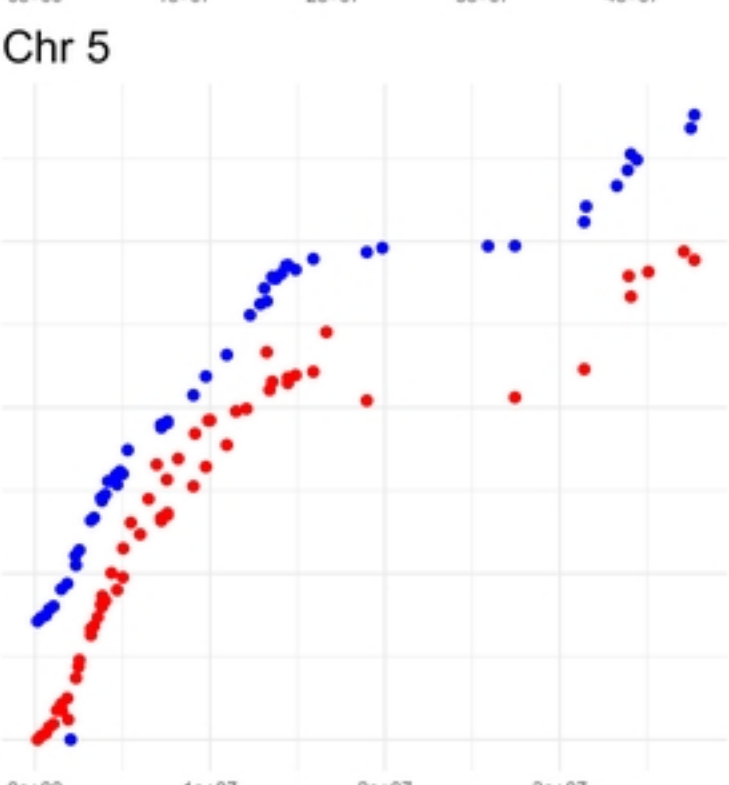
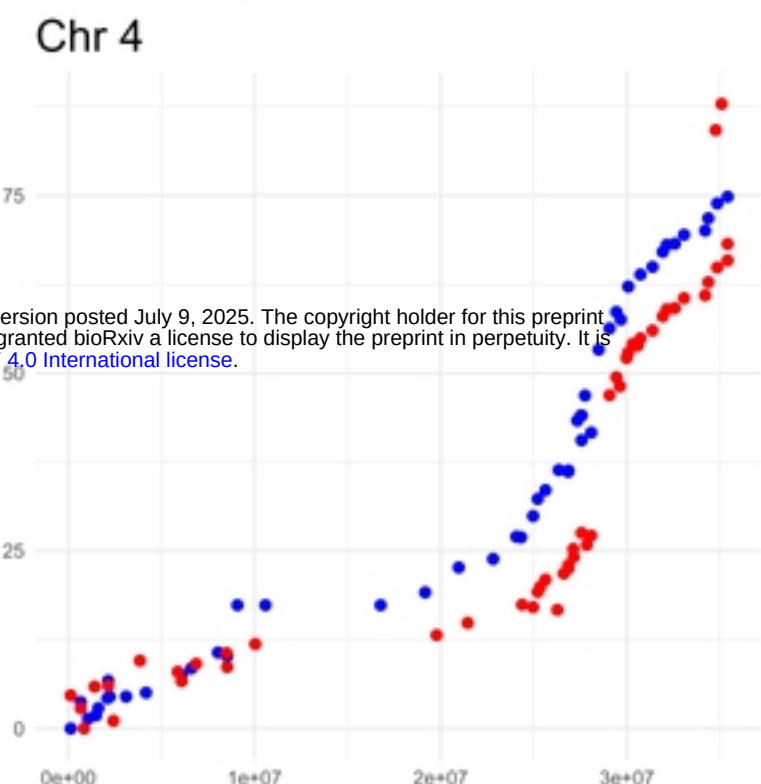
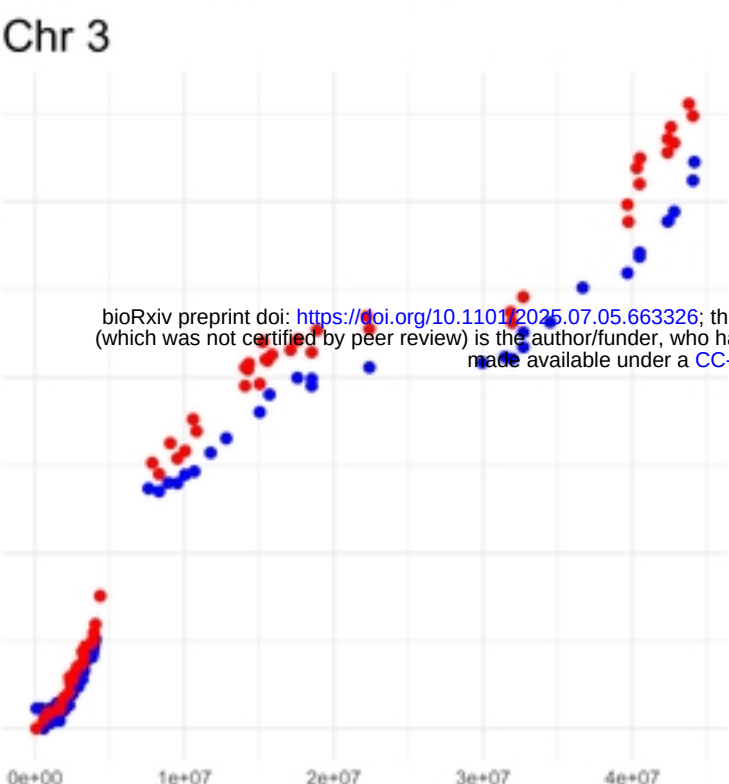
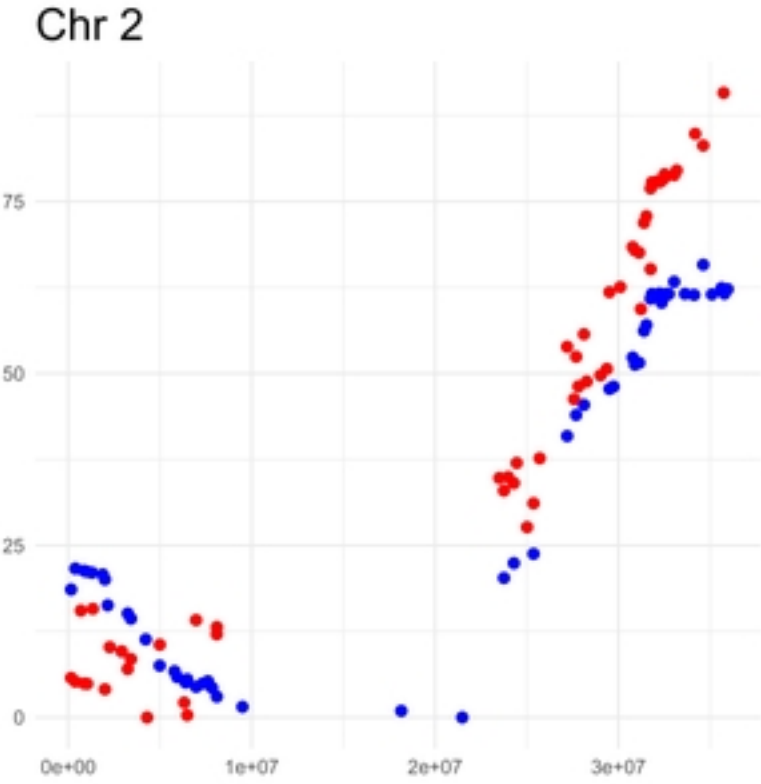
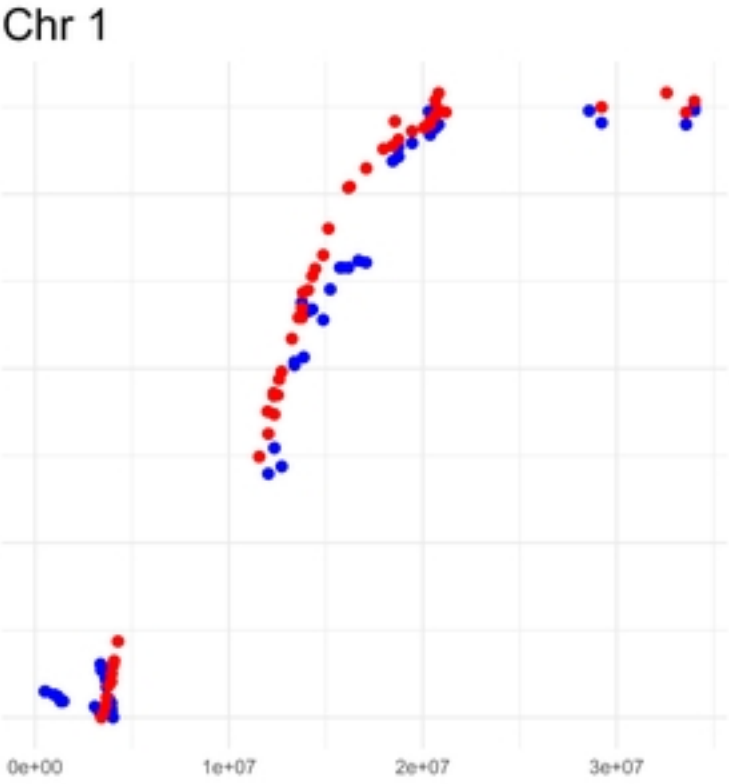


Figure2 Marker distribution



● Parent Paris
● Parent 486

x-axis labels = Physical position (bp)
y-axis labels = Genetic position (cM)

Figure3 Marey maps

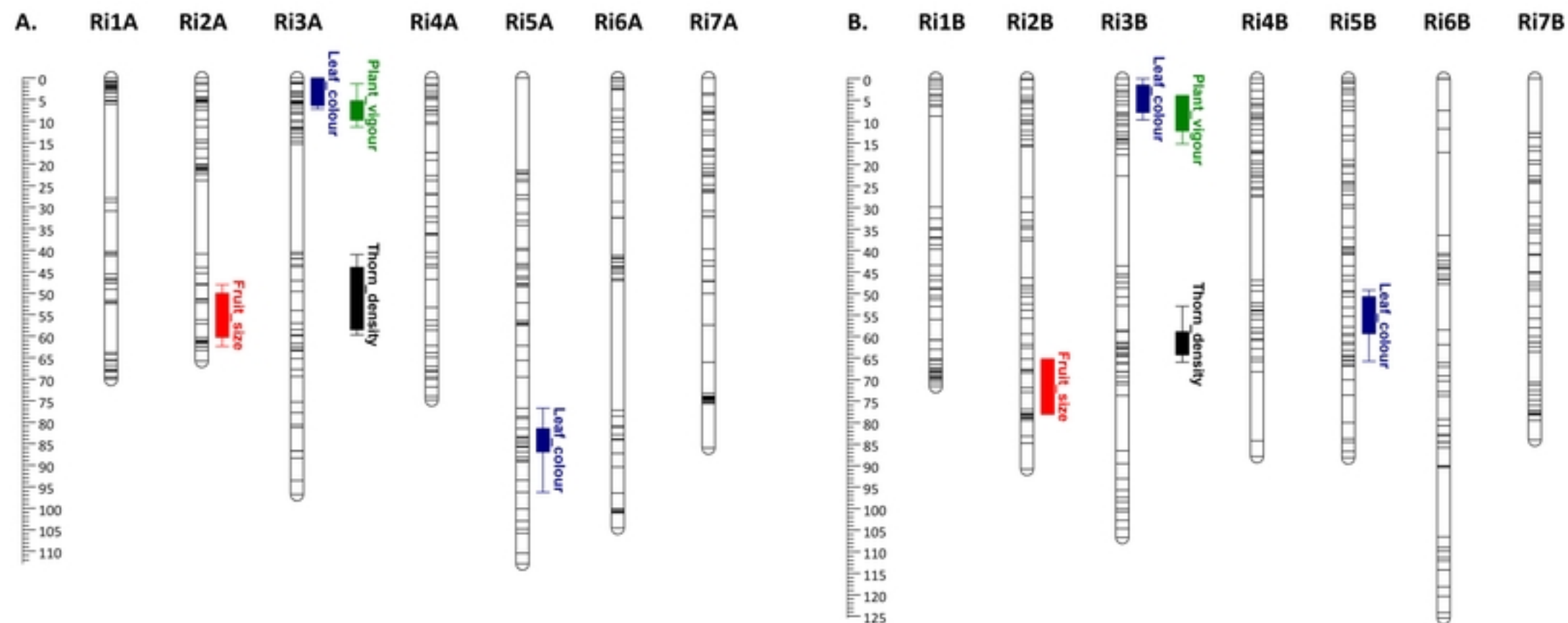


Figure4 QTL map

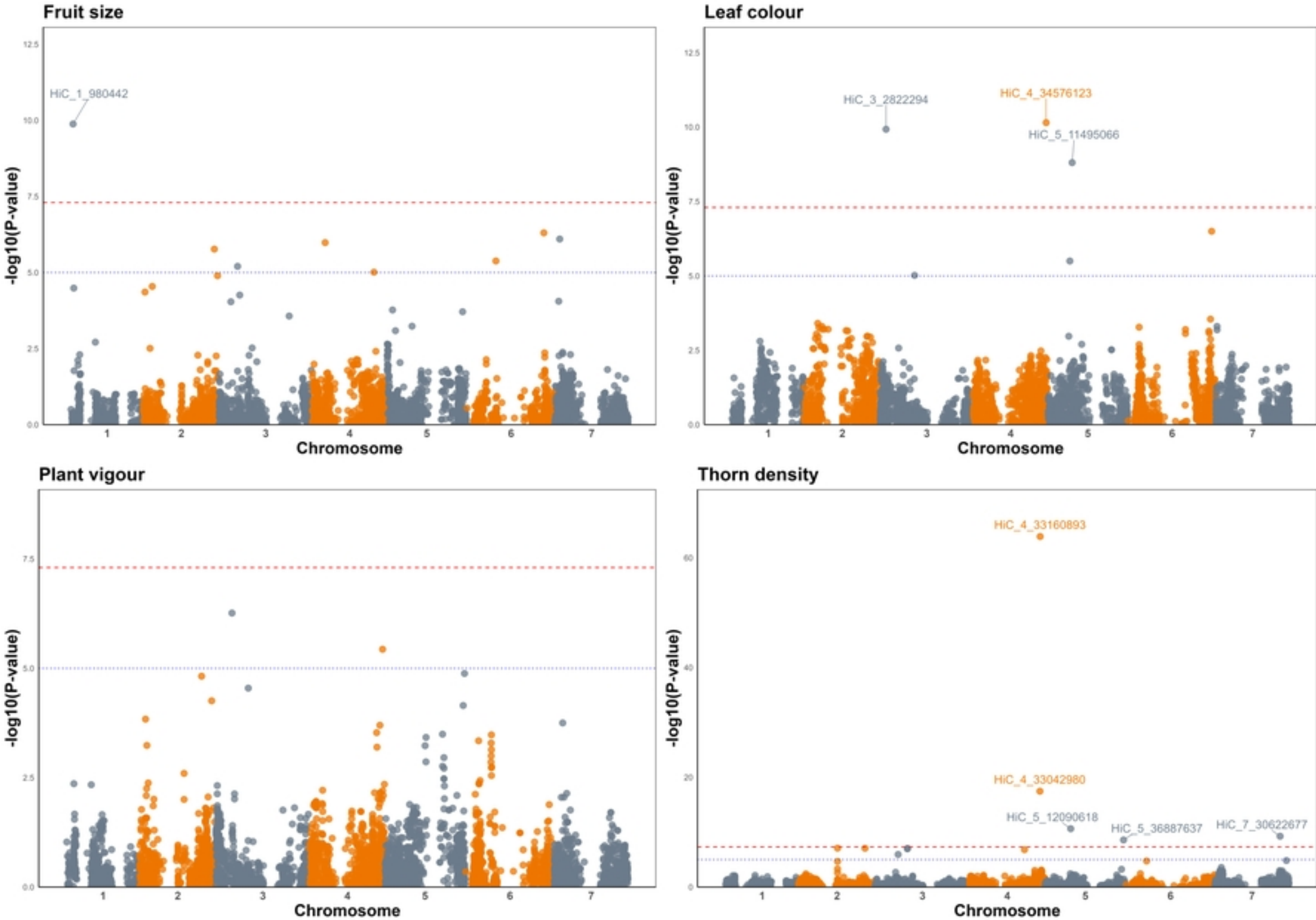


Figure5 Manhattan plots