The recombination landscape and multiple QTL mapping in a *Solanum tuberosum* cv. ‘Atlantic’-derived F1 population

Guilherme da Silva Pereira1,2\*, Marcelo Mollinari1,2, Mitchell J. Schumann1, Mark E. Clough1,3, Zhao-Bang Zeng1,2, G. Craig Yencho1\*

1Department of Horticultural Science, North Carolina State University, Raleigh, NC, USA

2Bioinformatics Research Center, North Carolina State University, Raleigh, NC, USA

3Vernon G. James Research and Extension Center, North Carolina State University, Plymouth, NC, USA

**\* Correspondence:**Guilherme da Silva Pereira  
gdasilv@ncsu.edu

G. Craig Yencho  
craig\_yencho@ncsu.edu

**Current addresses:**

Guilherme da Silva Pereira: Department of Genetics, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba, Brazil

Mitchell J. Schumann: Department of Soil and Crop Sciences, Texas A&M University, College Station, TX, USA

Keywords: tetraploid potato, linkage map, multiple interval mapping, mixed model, polyploid, yield, quality.

Abstract

There are many challenges involved with the genetic analyses of autopolyploid species, such as the tetraploid potato, *Solanum tuberosum* (). The development of a dosage-sensitive chip array and new analytical methods as well as newly collected data have made it valuable to re-analyze an F1 population () derived from a cross involving ‘Atlantic’, a widely grown chipping variety in the USA. A fully integrated genetic map with 4,486 single nucleotide polymorphisms (SNPs), spanning 1,629 cM, was constructed with MAPpoly software. We observed that bivalent configurations were the most abundant (51.0~71.7% depending on parent and linkage group), though multivalent configurations were also observed (2.1~40.3%). Seven traits were evaluated for four years (2006-8 and 2014) and quantitative trait loci (QTL) mapping was carried out using QTLpoly software. We detected 21 QTL for 15 out of 27 trait-year combination phenotypes. The most significant QTL was found on linkage group 5 for foliage maturity evaluated in 2008 and explained alone 64% of the phenotypic variation. This linkage group 5 region was identified as a QTL hotspot as other QTL co-located not only for maturity, but also for plant yield, specific gravity and internal heat necrosis (IHN) resistance. We found over 500 genes around QTL peaks, including those on chromosome 5 that have been previously implicated in maturity (*StCDF1*) and tuber formation (*POTH1*). These QTL regions can be investigated further in order to allow genomics-assisted breeding in tetraploid potato.

# Introduction

The potato is the third most important food crop in the world after rice and wheat in terms of human consumption. More than 19 million hectares were cultivated worldwide in 2017, with a production exceeding 388 million tons (FAO, 2020). Yet advancement in the breeding of new varieties is still relatively slow. ‘Russet Burbank’, which accounts for a majority of acreage planted in the United States, was developed in the early 1900s (Bethke et al., 2014), and ‘Atlantic’, a major chipping variety grown in the country, was originally released in 1976 (Webb et al., 1978). One of the reasons of such relatively slow breeding process is the relative complexity of the potato genome. Cultivated potato, *Solanum tuberosum*, is a highly heterozygous autotetraploid () outcrossing crop, which makes it very difficult to align potentially beneficial characteristic combinations through conventional breeding methods. In fact, many agronomic traits are polygenic, and show continuous distributions (Ghislain and Douches, 2020).

Much of the genome complexity have also impacted on the correct assessment of allelism from molecular markers, with consequences to linkage map construction and quantitative trait loci (QTL) mapping. Research has been conducted over the years to develop molecular techniques to facilitate potato breeding. However, since the first molecular genetic maps of potato were published (revised by Mann et al., 2011), very few breeding programs have begun to employ marker-assisted selection (MAS) routinely (Slater et al., 2014). Some of the exceptions to this are for more simple traits such as potato cyst nematode resistance (Schultz et al., 2012) and potato virus Y (Hämäläinen et al., 1997). More complex traits such as yield are still elusive, and much more work needs to be done.

The early linkage maps constructed for potato have incorporated qualitatively scored molecular markers (Mann et al., 2011). However, this traditional presence-absence based system limited not only the ability to explore the full range of segregating markers, but also to build integrated genetic maps for full-sib populations (Luo et al., 2004). Advances in molecular technologies have been made available to breeders with the publication of the diploid potato genome sequence (Xu et al., 2011) and the subsequent development and release of the Illumina Infinium® 8,303 Potato Array (Felcher et al., 2012) through the USDA-NIFA Solanaceae Coordinated Agricultural Project (SolCAP). This potato array is made up of 8,303 single nucleotide polymorphisms (SNPs) and has been shown to give very good coverage of the published genome potato (Felcher et al., 2012; Sharma et al., 2013). Utilizing Illumina’s technology of dual colored fluorescents to label the different nucleotides allows for allele dosage (Schmitz Carley et al., 2017; Zych et al., 2019). Along with the development of an array that uses dosage-sensitive SNPs, new analytical methods have recently been developed to effectively utilize this information in creating linkage maps (Bourke et al., 2018; Hackett et al., 2014; Mollinari and Garcia, 2019) and mapping QTL (Chen et al., 2018; Hackett et al., 2016) in autotetraploid species.

Novel linkage and QTL mapping approaches (da Silva Pereira et al., 2020; Mollinari et al., 2020) have allowed to explore new horizons when studying genetics of polyploid species. Particularly, Mollinari et al. (2020) presented a detailed characterization of the inheritance system in an even more complex autopolyploid species, the hexaploid sweetpotato, *Ipomoea batatas* (). As the inheritance pattern from parents to progeny had been unraveled, da Silva Pereira et al. (2020) were able to estimate identity-by-descent (IBD)-based additive relationship along the genetic map, which was ultimately used in a mixed model approach, facilitating multiple QTL mapping. In tetraploid potato, double reduction has been documented for several mapping populations (e.g. Bourke et al., 2015), but these studies lack a comprehensive assessment of multivalent formation and preferential chromosome pairing. In addition, single-QTL models used so far may have hindered the discovery of putative loci underlying the variation of agriculturally important traits in the species.

The advancement of molecular technologies and analytical methods, coupled with newly collected data, have made it valuable to re-analyze the B2721 potato mapping population, derived from ‘Atlantic’ cultivar (McCord et al., 2011b, 2011a; Schumann et al., 2017). We performed SNP dosage calling, created a fully phased, integrated linkage map, and carried out multiple QTL-based analyses for yield, foliage maturity, dry matter, specific gravity, and skin texture evaluated for four years. This has allowed us to infer the tetraploid inheritance mechanisms from a linkage analysis perspective and to better understand these traits at the genetic level as well as to bring breeders closer to MAS.

# Material and Methods

## Mapping population and field experiment

The cross ‘Atlantic’ × B1829-5 resulted in 156 progenies. This so-called B2721 mapping population has been chosen based on its segregation for internal heat necrosis (IHN), a non-pathogenic physiological disorder characterized by brownish spots in the tuber parenchyma (Yencho et al., 2008). ‘Atlantic’ is susceptible to IHN, whereas B1829-5, an advanced round white clone from the USDA-ARS Beltsville potato breeding program, is resistant. The population also segregates for other agronomic traits of interest, and these will be our focus in this work, although IHN related traits will also be considered. B2721 was previously evaluated in 2006, 2007 and 2008 (McCord et al., 2011a, 2011b) and additional data was collected in 2014.

The field designs for years 2006, 2007 and 2008 were described by McCord et al. (2011a, 2011b). Briefly, in 2006, the population was planted in an unreplicated trial with six plants per clone, whereas in 2007 and 2008, the population was planted in two replications with 10 plants per plot in a randomized complete block design, the same design also used in 2014. All experiments were carried out at the Tidewater Research Station (35°52'20" N, 76°39'33" W) in Plymouth, North Carolina, in a Portsmouth fine sandy loam. In 2014, untreated seed pieces were planted on March 21 and the harvest date was July 15 (a total of 116 days). Soil amendment and fertilizers were applied as follows: lime was applied on February 28, 363 kg of 15-15-15 fertilizer was applied on February 25, 76 liters of 30% nitrogen was applied on April 9, metribuzin and metalochlar were applied April 11, a tank mix of zeta-cypermethrin and beta-cyfluthrin was applied on May 12, 1 kg of manganese ethylenebisdithiocarbamate was applied on May 20, and imadacloprid was applied on June 3. At harvest, plots were dug using a chain digger and picked by hand. Tubers were then washed, culled for excessive rot, and phenotype measurements were obtained as described next.

## Phenotypic data and analyses

Phenotyping for all IHN-related and agronomic traits in 2006, 2007 and 2008 was described previously (McCord et al., 2011a). In 2014, in addition to IHN incidence (NI) and IHN severity (NS) (Schumann et al., 2017), the B2721 population was evaluated for five other traits: plant yield (PY), foliage maturity (FM), dry matter (DM), specific gravity (SG) and skin texture (ST). PY was measured as total plot weight minus the tubers culled divided by 10 (number of plants per plot) in kg/plant. FM was measured as the area under the senescence progress curve (Shaner and Finney, 1977), based on foliage ratings taken on June 9, June 25 and July 2, using a scale of 0 (0% yellowing) to 5 (100% senesced) with half point increments. For DM determinations, tubers were quartered from stem end to bud end using an Easy Wedger, Model N55550-4 (NEMCO Inc., Hicksville OH). For each plot, one quarter from four different tubers was placed into plastic whorlpak bags. Samples were then weighed, frozen at –20 °C, lyophilized, and weighed again. DM was calculated as the proportion between dry and fresh weights for each sample, in percentage. SG calculated using the formula by using the air weight () and water weight () for each plot. ST was rated on a 0 (smooth) to 3 (russeted) scale. Finally, NI was calculated as the proportion of tubers showing any sign of IHN over all evaluated tubers, and NS was rated on a 9 (no IHN signs) to 1 (completely necrotic) scale and averaged out for all evaluated tubers (Schumann et al., 2017).

Except for 2006 data (single observation), adjusted means were obtained separately for each trait-year combination based on a mixed model with genotypes and blocks as fixed and random effects, respectively, plus the random residual error using ASReml-R v. 4.1.0 (Butler et al., 2018). In addition, a joint analysis with data from 2007, 2008 and 2014 (replicated trials) was carried out, where years, genotypes and their interactions were treated as random effects. The resulting variance component estimates were then used to compute mean-basis broad-sense heritability (). Trait abbreviation followed by two-digit year was used to name each phenotype (trait-year combination, e.g. PY06). Progeny variation within each phenotype was explored with boxplots from R package ggplot2 (Wickham, 2016). Pearson’s correlations were computed using R v. 3.6.0 (R Core Team, 2019) and graphical visualization through network and correlation plots were obtained using R packages corrr v. 0.4.0 (Ruiz et al., 2019) and corrplot v. 0.84 (Wei et al., 2017). As an approach to investigate genotype-by-environment (GE) interaction, genotype (G) plus GE interaction (GGE) biplot analysis (Yan and Kang, 2003) was carried out using R package GGEBiplots v. 0.1.1 (Dumble et al., 2017) based on the standardized adjusted means.

## Genotypic data, dosage calling and linkage map construction

Genotyping was performed using the Illumina Infinium® 8,303 Potato Array (Felcher et al., 2012). The intensity of each fluorescent labeled nucleotide was read using the Illumina iScan Reader and imported into GenomeStudio. The ordered pair of normalized intensities from the two allelic variants was transformed into polar coordinates and to proceed with the dosage calling of the parents and offspring using the R package ClusterCall (Schmitz Carley et al., 2017). Uninformative markers or those resulting in unclassified parents were filtered out resulting in 5,599 SNPs.

Allele dosage from these markers were imported into MAPpoly v. 0.1.0 (Mollinari and Garcia, 2019) together with their respective chromosome positions (*S. tuberosum* genome v. 4.0.3). After testing for segregation distortion ( using Bonferroni correction) and filtering out redundant markers, the pairwise recombination fraction of the 4,812 SNPs (~11.6 million pairs) were computed for all possible linkage phases and the most likely configuration was selected. Markers were assembled into 12 linkage groups (LGs) using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) algorithm. Since 91.2% of markers within the groups coincide with their respective chromosomes, groups were formed using exclusively the genome information. Markers were ordered according the genome and phased using function ‘est\_rf\_hmm\_sequential’. Briefly, this function sequentially inserts markers into the map using two-point information to eliminate unlike linkage phases (LOD < 10.0), and for the remaining configurations, the multipoint LOD score is used (LOD < 10.0). If the insertion of the marker still results in less than 60 linkage phase configurations, the next marker is inserted and tested in the multiple configurations until the map is completed. In the final step, the multipoint recombination fraction was estimated considering a global genotyping error of 5%.

Next, the conditional probabilities of the 36 possible genotypes were obtained for every cM of the genome for all individuals in the offspring. These probabilities were used in further QTL analysis and to assess the meiotic process that formed the offspring. Probabilistic profiles for pairing behavior were obtained for all LGs for the three possible meiotic pairing configurations, i.e., *ab*/*cd*, *ac*/*bd*, and *ad*/*bc*, where *a*, *b*, *c* and *d* denote the homologs in parent ‘Atlantic’ and the notation *ab*/*cd* indicates that homolog *a* paired with *b*, and *c* paired with *d*, for example. The same reasoning applies to parent B1829-5 with homologs *e*, *f*, *g* and *h*. Also, the probabilistic haplotypes of all individuals in the offspring were reconstructed and the crossing-over points and respective homologs involved in the exchange were detected. Using this information and the heuristic algorithm presented in Mollinari et al. (2020), recombination chains were assembled for all individuals in the offspring, and the number and which homologs were involved in each meiosis were assessed. Recombination chains with more than two homologs involved imply that a multivalent formation was present during the meiosis.

## QTL mapping and gene search

Genotype conditional probabilities calculated from the genetic map were used to compute sib-pair IBD-based additive relationship matrix for every cM position. Using the R package QTLpoly v. 0.2.0 (da Silva Pereira et al., 2020), we selected putative QTL positions using a random-effect multiple interval mapping (REMIM) for each phenotype based on a forward-backward procedure. First, putative QTL were consecutively added to the model based on a forward search with a relatively relaxed genome-wide significance threshold (). Then, a backward elimination step was carried out under a more stringent threshold (). The variance components associated with the putative QTL were tested using linear score statistics (Qu et al., 2013), whose associated -values were log-transformed to facilitate QTL profile visualization with . The genome-wide significance was assessed using a score-based resampling method (Zou et al., 2004). A final multiple QTL model was fitted with the selected positions, and QTL genotype best linear unbiased predictions (BLUPs) were used to compute additive allele effects (Kempthorne, 1955). QTL heritability () was calculated as the ratio between the variance associated with the QTL and the total variance.

In order to compare different approaches, we also ran the fixed-effect interval mapping (FEIM) model (Hackett et al., 2014) as implemented in the QTLpoly package (da Silva Pereira et al., 2020). Using the same genotype conditional probabilities, a single-QTL model was used to fit six additive effects (as one additive effect from each parent is taken as reference) at every cM position. The model with a fitted QTL was compared to a null model (with no QTL) using likelihood ratio tests (LRT). LRT statistics were converted into LOD scores, and QTL were declared when the LOD score reached a threshold () based on 1,000 permutation tests (Churchill and Doerge, 1994). In both REMIM and FEIM analyses, a window size of 20 cM was used to avoid that two closely linked positions were declared as QTL. Approximate 95% QTL support intervals were estimated by dropping 1.5 from the QTL peak (REMIM) or (FEIM) (da Silva Pereira et al., 2020).

Based on the QTL peaks from REMIM, we searched for candidate genes in the region delimited by markers on the left and on the right of the QTL peak or within 200 kbp each side from the QTL peak, whichever was the largest, on the *S. tuberosum* v. 4.03 genome (ST4.03) using the PhytoMine tool (<https://phytozome.jgi.doe.gov/phytomine/>) by Phytozome v. 12 (Goodstein et al., 2012). Finally, we used the Web Gene Ontology Annotation Plot (WEGO) v. 2.0 (Ye et al., 2018; <http://wego.genomics.org.cn/>) to count and plot classifications of Gene Ontology (GO) terms for annotated genes within our QTL regions.

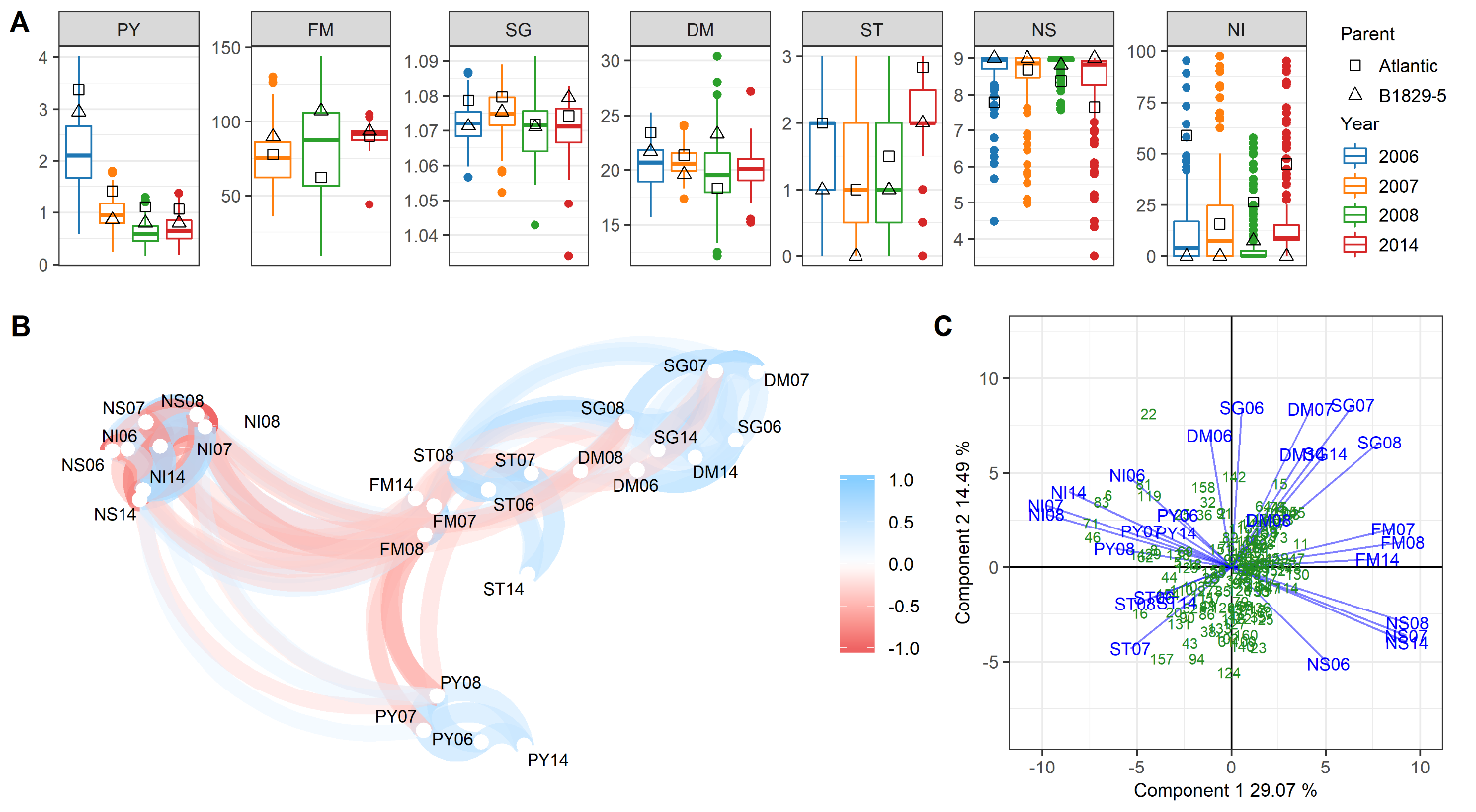
# Results

## Trait correlations and GE interactions

Based on the phenotype means (see Supplementary File S1), variation was found for all evaluated traits over the four years of evaluation (Figure 1A), indicating that they were segregating in the B2721 population. As the parental means were within the population phenotypic range, we could observe transgressive segregation for all trait-year combinations. PY06 has shown quite distinct median and broader variation in comparison to the following years, which is likely because there were no replicates in 2006, so that the environmental error could not be modeled. In contrast, FM14 exhibited a narrower variation when compared to the previous years, which was rather expected, since in 2014 maturity evaluation was carried out in a shorter time-window (24 days) in comparison to 2007 and 2008 (30 days). The overall change in ST14 scores when compared to the previous years was likely due to differences in evaluator scoring. Individuals seemed to be less affected by IHN in 2008, as NS08 and NI08 have shown distinct variation and median in comparison to the other years. Broad-sense heritability () ranged from 0.39 (DM) to 0.81 (SG), with similarly high values for PY (0.78), FM (0.70) and ST (0.79). For IHN-related traits, NS (0.74) showed a relatively higher heritability when compared to NI (0.66).

The correlation among variables was explored using a network plot (Figure 1B), where path color and transparency represent the signal and strength of the correlation, respectively. The variable positioning is given by multidimensional scaling that leverages the magnitude of the correlations, i.e. the closer the variables, the higher the absolute values of their correlations (Ruiz et al., 2019). In general, the phenotype of the same trait evaluated in different years clustered together. As expected, NS was close to NI as well as DM to SG, because each one of these trait pairs are different ways of assessing the same characteristic (IHN response and solids content, respectively). FM appeared centrally in this plot as it was strongly correlated with most variables either negatively (e.g. with PY and NI) or positively (e.g. with SG and DM). Note that for this network analysis, a subset of 144 full-sibs with complete observations among all traits was ultimately used. For correlation estimates based on pairwise complete observations, please see Supplementary Figure S1 and Supplementary File S2. Even though data from 2006 was originated from a single replication, they were in general agreement with the adjusted means from the other years. DM observations were the least correlated ones across years in comparison to all the other trait sets (see Supplementary Figure S1).

GGE biplot was used to investigate the existence of GE interaction (Figure 1C), where each year was considered an environment. The distribution of full-sibs along the first two principal components, which accounted for 43% of the variation, did not show any clustering. This is rather expected in mapping populations, usually consisting of non-selected individuals. However, some full-sibs were found to be more scattered on the fourth quadrant, indicating the most affected individuals by IHN (i.e. high NI). In contrast, more individuals were found to be less scattered on the second quadrant, where NS vectors lie, indicating smaller differentiation when full-sibs were less susceptible to IHN. Relatively shorter vector lengths of PY, ST and DM08 were indication of small capacity to discriminate genotypes using these traits, in comparison to longer vectors. Only relatively weak GE interaction was observed, since phenotypes from different years have consistently shown the same vector orientation for each trait, with small angles between vectors. Greater angles between vectors of phenotypes from 2006 (such as DM06, SG06 and NS06) and from the remaining years should be treated carefully, as these 2006 measurements were based on single plots. The same vector orientation for SG and DM indicated that these traits were positively correlated, whereas the opposite orientations for NS and NI vectors revealed their negative correlations, as already mentioned.

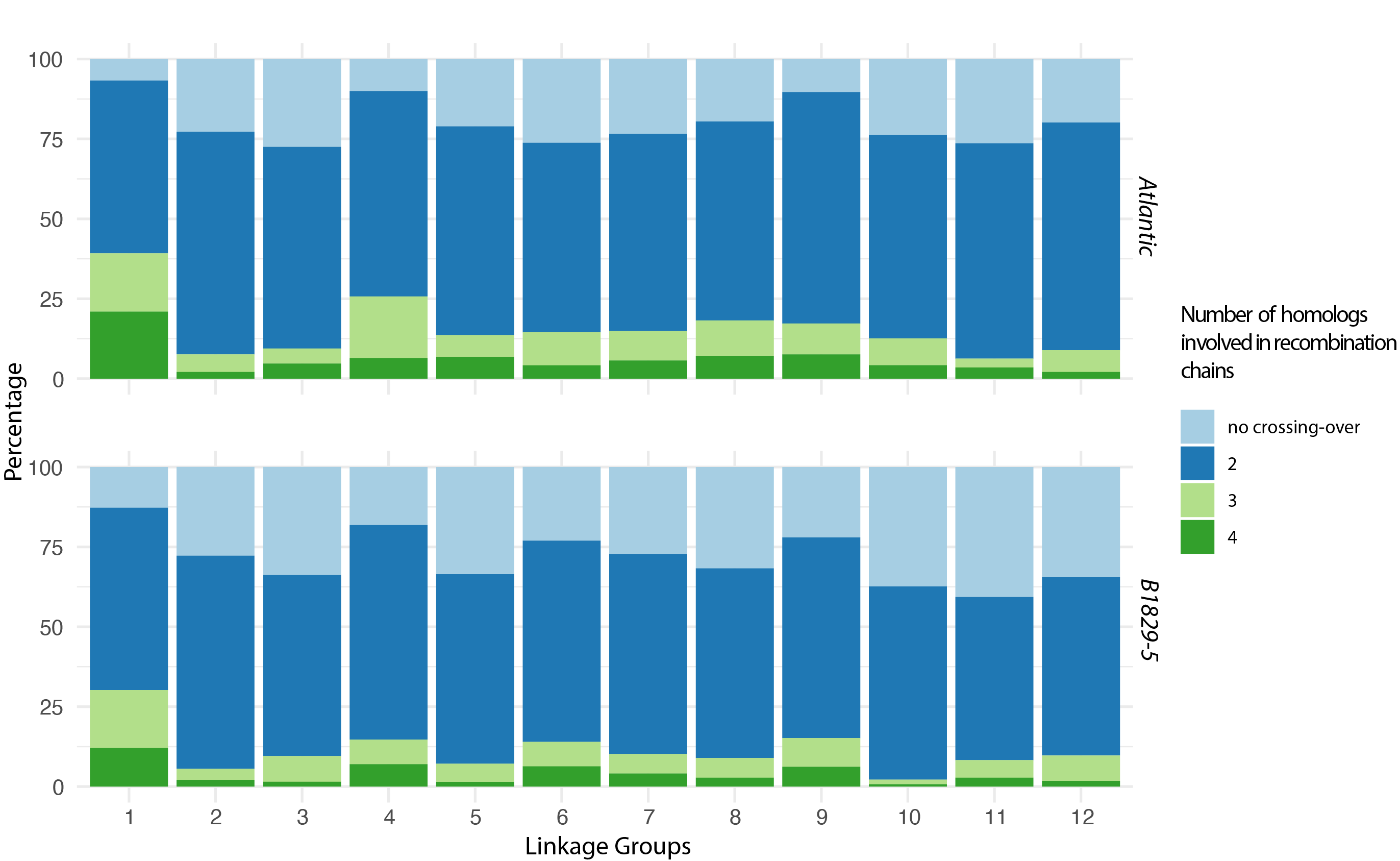


**Figure 1.** B2721 mapping population phenotypes evaluated for four years (2006-8 and 2014). (A) Boxplots showing distribution of full-sib means along with parental means. (B) Network plot showing significant correlations () in blue (positive) and red (negative). (C) GGE biplot showing the distribution of full-sibs (in green) and phenotypes (in blue) on the two first principal components. Traits: plant yield (PY), foliage maturity (FM), specific gravity (SG), dry matter (DM), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI).

## Linkage analysis and inheritance mechanisms

The normalized intensities from the Illumina Infinium® 8,303 Potato Array were obtained for the B2721 population (see Supplementary File S3). The dosage calling procedure yielded 5,599 informative SNPs with 1.05% of missing data. Fourteen percent of the SNPs were filtered out due to segregation distortion and redundancy resulting in 4,812 informative markers: 1,311 (27.3%) simplex, 1,234 (25.6%) double-simplex, and 2,267 (47.1%) multiplex (see Supplementary Figure S2). The complete linkage map (see Supplementary Table S1, Supplementary Figure S3 and Supplementary File S5) consisted of 12 LGs with eight haplotypes each, with four homologs per parent. A total of 4,285 markers spanned 1,629.99 cM in length, with an average density of 2.64 SNPs/cM. Linkage group lengths ranged from 106.20 cM (LG 5) to 205.88 cM (LG 1), with average LG length of 135.83 cM. The scatterplots of the physical distance in *S. tuberosum* genome v. 4.03 versus genetic distance in the 12 LGs in B2721 population map is shown in Supplementary Figure S4.

The probabilistic pairing profiles showed no preferential pairing between homologs in both parents (see Supplementary Figure S5). Among all meiotic configurations, only 7.3% and 9.8% were inconclusive for parents ‘Atlantic’ and B1829-5, respectively. From the remaining configurations (Figure 2), the percentage of cases with no crossing-over varied from 6.8% (LG 1, parent ‘Atlantic’) to 40.7% (LG 11, parent B1829-5), with mean 24.2%. Configuration involving two chromosomes with at least one crossing-over (i.e. bivalent configurations), were the most abundant varying from 51.0% (LG 11, parent B1829-5) to 72.4% (LG 9, parent ‘Atlantic’), with a mean of 62.3%. Multivalent configurations, i.e., involving three or four homologs, ranged from 2.2% (LG 8, parent B1829-5) to 39.2% (LG 1, parent ‘Atlantic’), with a mean of 13.5%.



**Figure 2.** Distribution of number of homologs involved in a recombination chain.

## Multiple QTL mapping and candidate genes

From the REMIM analysis, a total of 21 QTL were mapped in 15 out of 27 evaluated trait-year combinations (Table 1; see Supplementary Figure S6). Only one QTL was detected for 10 phenotypes each (for PY06, PY08, SG07, SG08, ST07, ST08, NS06, NS08, NI06 and NI08), and the other five phenotypes have shown two (for FM07, FM14, SG06 and NI07) or three (for FM08) QTL each. No QTL were mapped for the remaining 12 phenotypes (PY07, PY14, SG14, DM06, DM07, DM08, DM14, ST06, ST14, NS07, NS14 and NI14), mostly from years 2006 and 2014. Ten QTL were mapped on LG 5, four on LG 1, three on LG 7, and one on LGs 2, 3, 4 and 9 each. QTL heritability () ranged from 7.1% (FM08 on LG 7) to 64.0% (FM 08 on LG 5), including five QTL with high (), 10 with moderate (), and the remaining six QTL with lower () heritability (Table 1; Figure 3). In comparison to REMIM results, FEIM detected 17 QTL from 12 phenotypes (see Supplementary Figure S7). On the one hand, FEIM missed QTL previously identified for PY06 (on LG 5), FM08 (on LGs 1 and 7), SG06 (on LG 3), NI06 (on LG 1) and NI07 (on LGs 1 and 5). On the other hand, FEIM identified additional QTL for SG06 (on LGs 8 and 10), ST08 (LG 9), NI14 (on LG 5) (see Supplementary Table S1). The proportion of variance explained (PVE) by FEIM-derived QTL ranged from 13.0% to 52.7%.

**Table 1.** Random-effect multiple interval mapping (REMIM) for B2721 traits evaluated for four years (2006-8 and 2014).

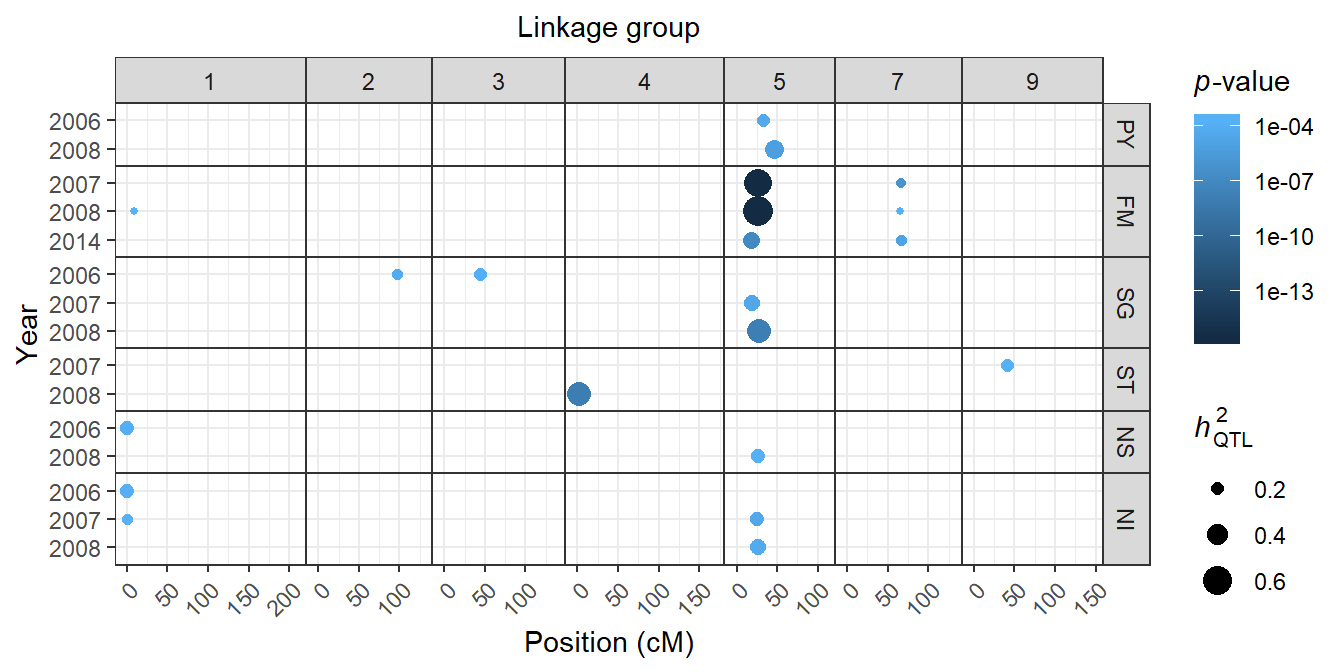
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Traita | QTL | Linkage  group | Position  (cM) | SI (cM)c | Score | -value | d | e  (%) |
| PY06 | 1 | 5 | 32 | 13-42 | 137.41 | 4.70e–05 | 1.11e–01 | 20.4 |
| PY08 | 1 | 5 | 46 | 0-53 | 168.95 | 7.21e–06 | 2.17e–02 | 36.0 |
| FM07 | 1 | 5 | 26 | 18-29 | 355.99 | <2.22e–16 | 3.36e+02 | 57.7 |
|  | 2 | 7 | 66 | 52-70 | 167.43 | 8.98e–07 | 7.28e+01 | 12.5 |
| FM08 | 1 | 1 | 9 | 0-32 | 101.58 | 2.16e–04 | 1.17e+02 | 7.4 |
|  | 2 | 5 | 27 | 0-41 | 425.04 | <2.22e–16 | 1.01e+03 | 64.0 |
|  | 3 | 7 | 65 | 35-71 | 104.85 | 2.00e–04 | 1.13e+02 | 7.1 |
| FM14 | 1 | 5 | 18 | 0-29 | 197.01 | 9.77e–08 | 7.12e+00 | 29.8 |
|  | 2 | 7 | 66 | 54-70 | 138.45 | 1.57e–05 | 4.01e+00 | 16.8 |
| SG06 | 1 | 2 | 97 | 34-117 | 125.63 | 6.65e–05 | 6.69e–06 | 15.9 |
|  | 2 | 3 | 44 | 31-50 | 109.89 | 1.49e–04 | 8.94e–06 | 21.2 |
| SG07 | 1 | 5 | 19 | 4-32 | 135.93 | 3.26e–05 | 1.39e–05 | 27.6 |
| SG08 | 1 | 5 | 28 | 15-32 | 240.07 | 1.47e–08 | 4.24e–05 | 48.9 |
| ST07 | 1 | 9 | 41 | 27-82 | 109.68 | 1.97e–04 | 1.59e–01 | 21.9 |
| ST08 | 1 | 4 | 2 | 0-4 | 234.05 | 1.08e–08 | 6.32e–01 | 48.1 |
| NS06 | 1 | 1 | 0 | 0-52 | 113.07 | 1.29e–04 | 1.17e–01 | 22.9 |
| NS08 | 1 | 5 | 26 | 12-32 | 117.77 | 1.77e–04 | 2.84e–02 | 23.4 |
| NI06 | 1 | 1 | 0 | 0-69 | 110.07 | 1.66e–04 | 9.28e+01 | 22.0 |
| NI07 | 1 | 1 | 0 | 0-8 | 103.32 | 2.26e–04 | 1.39e+02 | 16.8 |
|  | 2 | 5 | 25 | 8-34 | 134.43 | 3.35e–05 | 1.99e+02 | 24.2 |
| NI08 | 1 | 5 | 27 | 12-32 | 131.28 | 6.54e–05 | 5.65e+01 | 28.5 |

aTraits: plant yield (PY), foliage maturity (FM), specific gravity (SG), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI).

b~95% support interval (SI) based on , in centiMorgans.

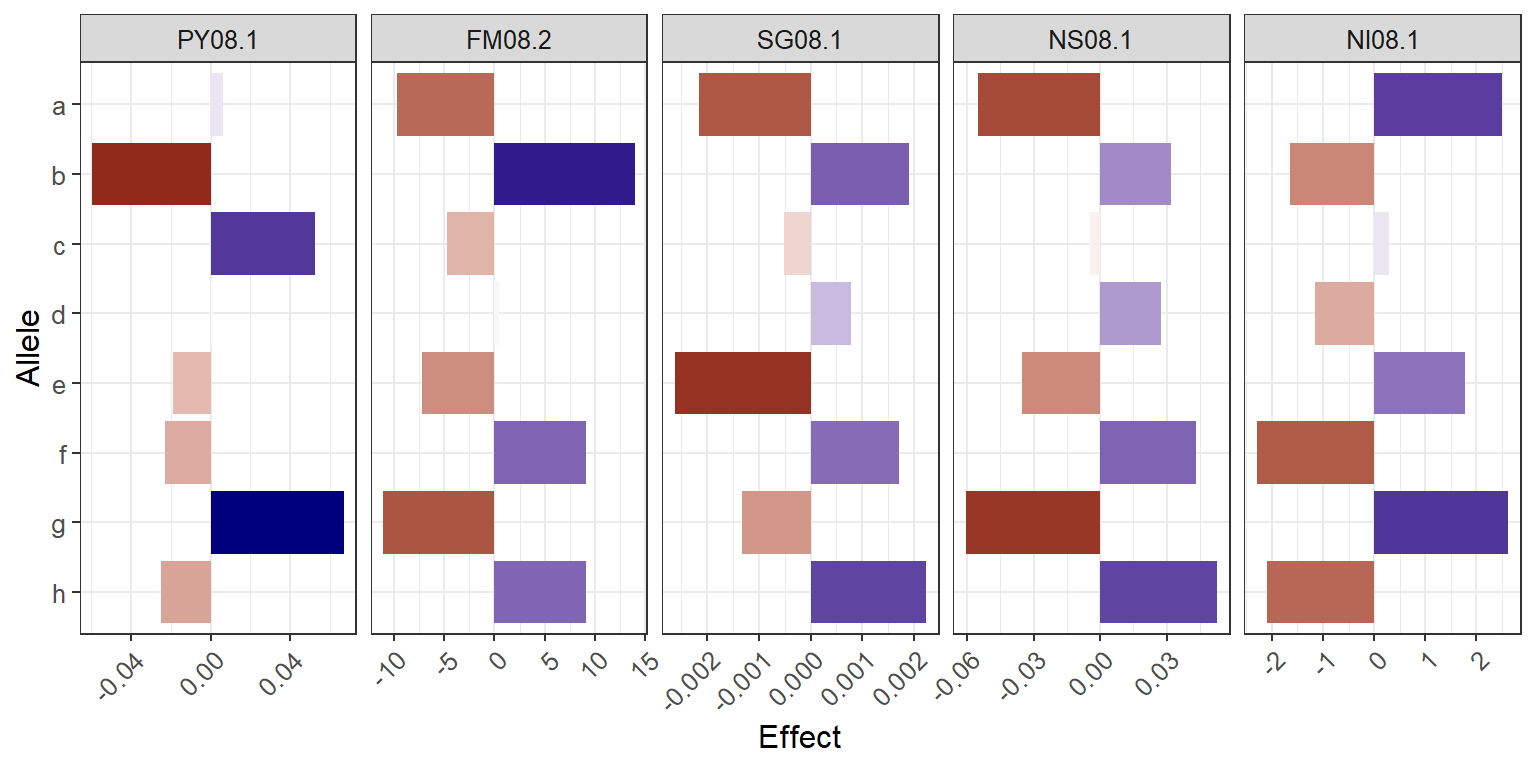
cVariance component associated with the QTL ().

dQTL heritability (), in percentage.



**Figure 3.** QTL identified for six traits evaluated for four years (2006-8 and 2014). Dots are plotted according to QTL peak location along the linkage groups. Color scale represents -values, and sizes are proportional to QTL heritability (). Traits: plant yield (PY), foliage maturity (FM), specific gravity (SG), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI).

With 10 QTL on the proximal half of LG 5, this region was characterized as a QTL hotspot. The major QTL on LG 5 for FM was detected for all three years of evaluation (2007-8 and 2014). However, profiles (see Supplementary Figure S6) showed that there were potentially more QTL underlying the variation of traits where no QTL was declared. For instance, suggestive QTL (at a lower that did not reach the threshold) can be seen on LG 5 for PY06, SG14 and NS14 and NI14. For the five traits with QTL on LG 5 (PY, FM, SG, NS and NI), most QTL peaks were found at 29 cM and support intervals ranged from 0 to 50 cM. Figure 4 depicts additive allele effects of co-located QTL on LG 5 for these different traits evaluated in 2008. It is interesting to note that for the QTL sets PY08.1/NI08.1 and FM08.2/SG08.1/NS08.1, major allele effects appeared in the same direction within sets, but in opposite directions between them. As these QTL on LG 5 were responsible for explaining great portion of the variation for these traits (Table 1), the QTL-based predicted means were highly correlated, either positively (PY with NI and FM with SG and NS) or negatively (PY/NI with FM/SG/NS) (see Supplementary Figure S8). It is worth to mention that selecting towards two out of four alleles per parent will tend to favor different sets of phenotypes depending on the alleles. For example, selecting allele pair *ac* from ‘Atlantic’ and *eg* from B1829-5 will likely increase PY and NI and decrease FM, SG and NS.



**Figure 4.** Additive allele effects for co-located QTL on LG 5 for five traits evaluated in 2008. Letters represent each parental haplotype from the linkage map (‘Atlantic’ = *a* through *d*, B1829-5 = *e* through *h*). Traits: plant yield (PY), foliage maturity (FM), specific gravity (SG), and internal heat necrosis severity (NS) and incidence (NI).

As the QTL support intervals comprised too large of a region to manually curate (>13 Mb per QTL, on average), we focused our candidate gene search on the region within the markers flanking the QTL peak (~400 kb, on average) (see Supplementary Table S2). We have learned that 533 annotated genes on *S. tuberosum* v. 4.03 genome were within these QTL regions, and that 342 genes had been assigned to at least one GO term (see Supplementary File S6). From a total of 594 GO terms, 309 referred to a different set of molecular function (mostly related to catalytic activity or binding), 208 to a biological process (greatly represented by metabolic or cellular processes), and 77 to a cellular component (such as membrane or another cell part) (see Supplementary Figure S9).

1. **Discussion**

For the same B2721 population, two separate linkage maps, one for each parent, have been previously built using amplified length polymorphisms (AFLP) and a few simple sequence repeat (SSR) markers, allowing varied degrees of map integration within parents (McCord et al., 2011a). That is, not all haplotypes had been identified based on the marker technology and analytical methods available at the time. Another issue was anchoring LGs to chromosomes. Despite the use of a few SSRs to anchor eight out the 12 LGs to *S. tuberosum* chromosomes (McCord et al., 2011a), QTL comparisons were limited between studies due to the fact that AFLP are not informative regarding physical mapping. Schumann et al. (2017) have already shown how the dosage-sensitive SNP array contributed to map integration and haplotyping in the B2721 population. However, their map contained 3,427 SNPs and spanned 1,397.86 cM, which is relatively smaller but less saturated than our current map with 4,486 SNPs distributed along 1,628.77 cM.

A significant difference in our analysis was the usage of the multipoint-based algorithm implemented in MAPpoly. Schumann et al. (2017) used the software TetraploidSNPMap (Hackett et al., 2016), which relies on simplex markers to assemble homologs, bridging these homologs afterwards using multidose markers. Thus, the resulting map depends heavily on the distribution of simplex markers along each homolog. On the other hand, MAPpoly algorithm considers the four homologs for each parent at the very beginning of the map construction and inserts markers sequentially regardless of their dosage. Since more than 72% of the markers were double-simplex and multiplex, our mapping strategy was more suitable to analyze the B2721 population. Also, we used the *S. tuberosum* v. 4.0.3 genome order instead of relying exclusively on genetic distance to obtain the marker order, e.g., using the Multidimensional Scaling (MDS) algorithm (Preedy and Hackett, 2016). Therefore, our genome-assisted approach helped to include markers that otherwise would be excluded using only two-point based linkage analysis.

As expected, the probabilistic pairing profiles showed no preferential pairing between homologs in both parents (see Supplementary Figure S5), likely due to the cultivated potato’s autotetraploid nature. Consequently, alleles present in different homologs in the same homology group had an equal chance to recombine with alleles from other loci, amplifying the range of possible genotypes when compared to diploids or allotetraploids. In addition, most of the parental meiotic configurations based on the B2721 population (62.3%, in average) were inferred as bivalents since they involved two homologs exchanging segments during the metaphase I (Figure 2). These results are in accordance with recent findings published by Choudhary et al. (2020). These authors presented a map of meiotic stages in *S. tuberosum* in one diploid and three tetraploid potato varieties using fluorescence *in situ* hybridization (FISH). They concluded that bivalent chromosome associations are the most common in metaphase I of tetraploid potatoes. Likewise, the broad range of multivalent signature rates observed in our study (2.2~39.2%) was also observed by Choudhary et al. (2020) (7~48%). Similarly, as their observations differed depending on variety and chromosome, in our study, the variation occurred by parents and by linkage groups. Thus, our meiotic assessment, although not as precise as a cytological analysis and prone to sampling errors, can be very helpful, serving as a proxy to evaluate meiotic configurations by using a straightforward expansion of linkage analysis.

Previous QTL mapping analyses in the B2721 mapping population were based on an interval mapping model, and the 95th percentile () of maximum LOD scores from 1,000 permutation tests per phenotype was used to declare QTL, with several additional suggestive QTL (below this threshold) also being recorded (McCord et al., 2011a, 2011b). It is worth to mention that the lack of integration within and between parental maps may have resulted in the same region being declared as QTL more than once. In this case, the QTL alleles represented in separate parental haplotypes might still contribute significatively to the phenotypic variance, when compared to the haplotypes with alternate alleles. However, this is a rather than desired outcome, since geneticist’s expectations are that one can learn how the whole haplotypic set contributes to the variation of phenotypic traits in the population, instead of being limited to the separate parental haplotypes.

For plant yield, McCord et al. (2011a) have listed a total of seven QTL across three years, PY06 (one), PY07 (five) and PY08 (two), from both parental maps, in addition to 11 suggestive ones. On the other hand, current REMIM analyses resulted in a single co-located QTL region on LG 5 for PY06 and PY08 (Table 1; Figure 3). QTL landscapes for PY07 and PY14 have shown suggestive QTL at the same location (see Supplementary Figure S6), but lower significances did not allow neither REMIM nor FEIM to define them as QTL. In fact, FEIM has even missed the QTL for PY06 (see Supplementary Figure S7). QTL on LG 5 for tuber yield have also been reported in both diploid (DM 1-3 516 R44 × RH89-039-16, ) (Manrique-Carpintero et al., 2015) and tetraploid (‘Liberator’ × W4013-1, ) (Rak et al., 2017) potato mapping populations.

For foliage maturity, McCord et al. (2011a) and Schumann et al. (2017) have described major QTL for years 2007-8 on LG 5 as well as the one on LG 7. In the current study, another QTL on LG 1 appeared for FM08. Despite the differences regarding phenotypic distribution for FM14 (Figure 1A), we were able to find QTL on LGs 5 and 7 (Figure 3). This LG 5 region is well known as a main source of QTL in potato. Prior to the release of the SNP potato array, a major QTL for plant maturity had already been reported on chromosome 5, also linked to a late blight (*Phytophthora infestans*) resistance locus (reviewed by Danan et al., 2011). More recently, Massa et al. (2018) noticed that the same region on LG 5 also underly the variation of several tuber quality-related traits in addition to resistance to early blight (*Altenaria tenuis*) and Verticillium wilt (*Verticillium* spp.) for a russet mapping population (‘Rio Grande Russet’ × ‘Premier Russet’, ). In fact, For IHN-related traits, three QTL were identified on LG 5, one for each NS08, NI07 and NI08, using REMIM. FEIM identified an additional significant QTL on LG 5 for NI14 (see Supplementary Table S3 and Supplementary Figure S7). Previously, IHN-related QTL had been reported as suggestive only (Schumann et al., 2017), and this difference in significance is likely due to an improved map construction.

Additionally, a major QTL on LG 5 was found to co-locate for SG07 and SG08, while it appeared only as suggestive for SG14 and with no evidence for SG06. However, there was no evidence for QTL for dry matter. Specific gravity and dry matter are supposedly related with solid content in the tuber. In fact, the correlation among their adjusted means were generally high (up to 0.81 between DM07 and SG07; see Supplementary Figure S8 and Supplementary File S2), but the lack of co-located QTL limited any conclusions regarding any shared genetic control. It is worth to mention that their broad-sense heritability values () differed greatly (0.39 for DM and 0.81 for SG), so that the ability to detect QTL for DM may have been affected because smaller proportion of the total variance could be attributed to genetic variance. Other mapping populations have revealed QTL on LG 5 for DM (12601ab1 × ‘Stirling’, ) (Bradshaw et al., 2008) and SG (Manrique-Carpintero et al., 2015), where somewhat distinct heritability values were found (0.81 for DM and 0.46 for SG).

Skin texture was a trait which diverged the most from the other traits regarding QTL locations, specifically for not showing enough statistical evidence for QTL on LG 5 (see Supplementary Figure S6). In diploid potatoes, skin texture is believed to be controlled by three loci (Jong, 1981). In our B2721 population, we detected two QTL (one for ST07 and ST08 each), in regions on LGs 4 and 9 where no other traits have shown evidence for QTL in our study. That is, the apparent strong correlation observed between ST07 and FM07 of –0.44 (), for example, could not be attributed to genetics, as correlation between QTL-based predicted means for these two traits was –‍0.07 (; see Supplementary Figure S8).

Two main reasons may have contributed to the relative low QTL discovery in this population. First, cultivars and elite clones, such as those used as B2721 parents, have usually been selected such that alleles for characteristic of interest were led to fixation. For instance, as observed by Schumann et al. (2017), IHN-related traits were concentrated towards resistance (Figure 1A). It is likely that one or a few major loci underlying IHN resistance have their alleles fixed within a parent (B1829-5) or, at least, that these allele effects do not differ greatly, hence their contribution towards resistance is more or less the same. For the QTL on LG 5 for NI08, the allele contributions to the population mean (~6.1%) were ±2.5% at most (Figure 4). In other words, these effects would not change the fact that most individuals would be less susceptible (NI < 10%) because resistance could have been already conferred by other undetectable QTL. In F1 populations, the fact that two parental genotypes are contrasting does not directly imply that major QTL will be detected. For instance, considering a biallelic QTL, a cross between two contrasting parents, *QQQQ* × *qqqq*, would not segregate (100% *QQqq*). In order to be detectable, alleles need to be contrasting within parent (da Silva Pereira et al., 2020), e.g. *Qqqq* × *qqqq* or *Qqqq* × *Qqqq* in the simplest cases, resulting in segregations such as 50% *Qqqq* : 50% *qqqq* or 25% *QQqq* : 50% *Qqqq* : 25% *qqqq*, respectively.

Another reason for limited detection power is, if there were alleles with contrasting effects within parents, small population sizes may not allow to separate signal from noise properly. Consequently, thresholds for declaring a QTL are usually required to be more stringent in order to avoid false positives. In general, tetraploid potato QTL mapping studies have been performed in relatively small population sizes (~150 or less individuals). In these cases, even high significances and QTL heritability estimates, such as those found for FM07, FM08 and SG08, should be treated carefully as they can be relatively biased due to sampling. Finally, reduced population sizes also limit the ability to test closely linked versus pleiotropy for co-located QTL across different traits accurately.

Among the candidate genes (see Supplementary File S6), GO terms that overlapped DNA and protein binding functions with regulation of biological process were of particular interest. In total, 19 genes were found to have the same two GO terms (GO:0003677 and GO:0006355) attributed to the *S. tuberosum* CYCLING DOF (DNA-binding with one finger) FACTOR 1, *StCDF1* (PGSC0003DMG400018408, ST4.03ch05:4539029..4541329), which has been described as the transcription factor underlying the major QTL for maturity on LG 5 (Kloosterman et al., 2013). Such a region also includes yield-related QTL, where we found the *Potato homeobox 1* gene, *POTH1* (PGSC0003DMG400013493, ST4.03ch05:9248372..9258054), that appears to regulate tuber development in potato via StBEL5-TOPH1 heterodimer (Rosin et al., 2003). Finally, QTL underlying IHN-related traits were also found in the same LG 5 hotspot, where a homeobox-associated leucine zipper, *StHOX40* (PGSC0003DMG400030494, ST4.03ch05:4264626..4267698), was identified. This gene was found to be either up or downregulated under heat and drought stresses in potato, respectively (Li et al., 2019), which are believed to be two important triggers of IHN (Yencho et al., 2008).

Other genes, such as *StERF6* (PGSC0003DMG400016651, ST4.03ch06:33827698..33832740) and *StNAC046* (PGSC0003DMG400031266, ST4.03ch05:4946232..4948779), were characterized as transcription factors from AP2/ERF and NAC families, respectively, which have been implicated in a wide range of regulation processes in plants (Mizoi et al., 2012; Olsen et al., 2005). A gene from the NAC family, *StNAC032* (PGSC0003DMG400002824, ST4.03ch04:245125..246607), was retrieved from the LG 4 region of a major QTL for skin texture. Several genes from the NAC family, including *StNAC032*, were found to be expressed in the tuber skin and appeared to be involved in suberin and associated wax biosynthesis (Soler et al., 2020), which could contribute towards skin texture, as well as be involved with response to drought and other biotic and abiotic stresses (Singh et al., 2013).

In plants, a single QTL region may be found to underly the variation of more than one phenotype at once, which can potentially explain the correlation among traits. A QTL hotspot has been identified in the short arm of chromosome 5 of potato (D’hoop et al., 2014; Danan et al., 2011), and our B2721 population stuck to the rule. A major QTL for maturity was found to overlap loci underlying the variation for other traits such as yield, specific gravity and IHN-related resistance, as described previously (McCord et al., 2011b, 2011a; Schumann et al., 2017) and corroborated with newly analyzed phenotypic data. Pleiotropy is one possible explanation of physiologically correlated traits, such as maturity, yield and specific gravity (as a proxy of solid content, thus contributing to yield). Transcription factors are known to have this property in which a single molecular function may be involved in multiple biological processes (He and Zhang, 2006). For instance, a DNA binding element from the HAP complex, *Ghd8*, was found to play pleiotropic roles in regulating yield, flowering and plant height in rice (Yan et al., 2011). On the other hand, the presence of QTL for yield-related traits and pathogenic (e.g. late blight) and non-pathogenic (e.g. IHN) potato disorders in the same hotspot on LG 5 could well be due to closely linked genes. Another example from rice has suggested the multigenic nature of a major QTL for grain yield under drought, where the central role of a NAM transcription factor, *OsNAM12.1*, in concert with extra co-localized genes, influenced other traits such as root and panicle branching, and transpiration efficiency (Dixit et al., 2015). In potato, our list of candidate genes can be used in further experiments to dissect such a QTL hotspot. Importantly, genomics-assisted breeding for quantitative, more complex traits could be implemented upon validation of these QTL neighboring markers in breeding populations of tetraploid potato.

# Author Contributions

GCY designed the experiments and supervised the project. MJS and MEC performed field experiments and collected data. GSP, MM and ZBZ analyzed data. GSP drafted the manuscript with major contributions from MM. All authors reviewed and approved the manuscript.

# Funding

This project was funded by USDA-ARS [58-1245-3-307, 2014-34141-22266] and USDA-NIFA [2013-34141-21392]. Scholarships were awarded to MJS by Pioneer and Monsanto, and to GSP and MM by Bill & Melinda Gates Foundation [OPP1052983]. Currently, GSP holds a scholarship from PrInt-CAPES, Brazil [88887.369781/2019-00].

# Acknowledgments

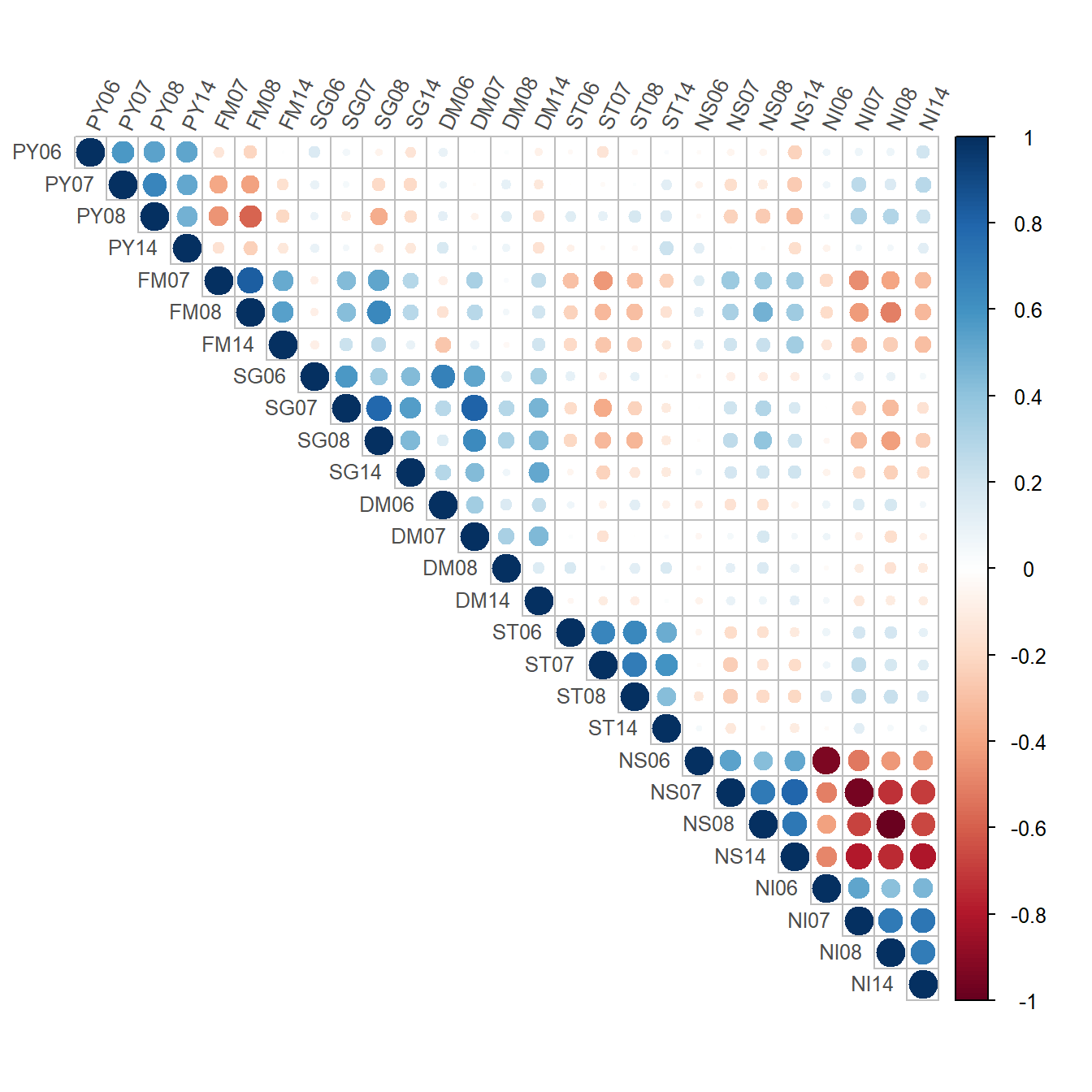
We would like to thank all the staff at the North Carolina Department of Agriculture and Consumer Services Tidewater Research Station, Plymouth, NC, for their hard work and dedication to quality research in aiding with the field aspect of this study. We also appreciate Michigan State University Potato Breeding and Genetics Program for their assistance with the Illumina Infinium® 8,303 Potato Array.

# Data Availability Statement

The datasets analyzed in this study can be found in the supplementary material. Both MAPpoly (<https://github.com/mmollina/MAPpoly>) and QTLpoly (<https://github.com/guilherme-pereira/QTLpoly>) R packages are available at GitHub.

# Supplementary Material

## Supplementary Figures



**Supplementary Figure S1.** Correlation between adjusted means for B2721 traits evaluated for four years (2006-8 and 2014) based on pairwise complete observations. Traits: plant yield (PY), foliage maturity (FM), specific gravity (SG), dry matter (DM), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI). (PDF)

**A picture containing curtain

Description automatically generated**

**Supplementary Figure S2.** Filtered data used to build the B2721 population map: 4812 SNPs scored in 156 individuals. The barplot indicates the dosage combinations for all markers: 1-0 indicates 0 doses for parent ‘Atlantic’ and 1 dose for parent B1829-5, and so on. The blue dots indicate the for a chi-square test of segregation distortion under Mendelian inheritance. The colored panel indicates the distribution of the dosages in the population.

**A screenshot of a cell phone

Description automatically generated**

**Supplementary Figure S3.** Genetic map of the B2721 population, with black vertical lines representing SNPs in their respective positions. (PDF)

**A close up of a map

Description automatically generated**

**Supplementary Figure S4.** Scatterplots of the B2721 genetic map (in centiMorgans, cM) versus *Solanum tuberosum* v. 4.03 reference genome (in mega base pairs, Mbp). (PDF)

A screenshot of a cell phone

Description automatically generated

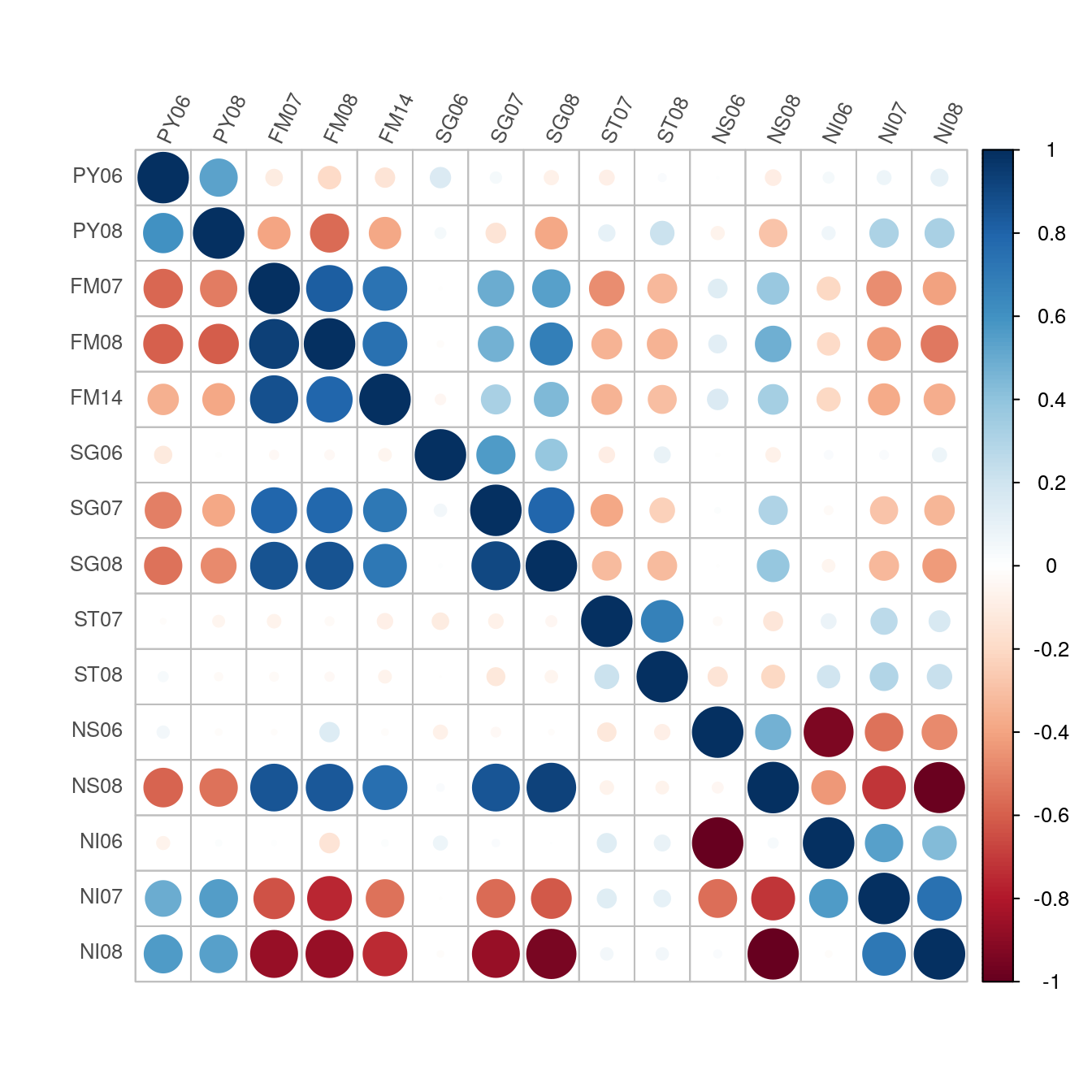
**Supplementary Figure S5.** Homolog pairing assessment in B2721 mapping population. (A) Probabilistic pairing profiles, where parental homologs (‘Atlantic’ = *a* through *d*, B1829-5 = *e* through *f*) are paired according to the following notation: e.g. *ab*/*cd* where homolog *a* paired with *b*, and *c* paired with *d*. The dashed line is the expected probability under random pairing (1/3). (B) -values associated to the chi-square test with the null hypothesis that all pairing configurations have the same probability. The dashed line represents .

****

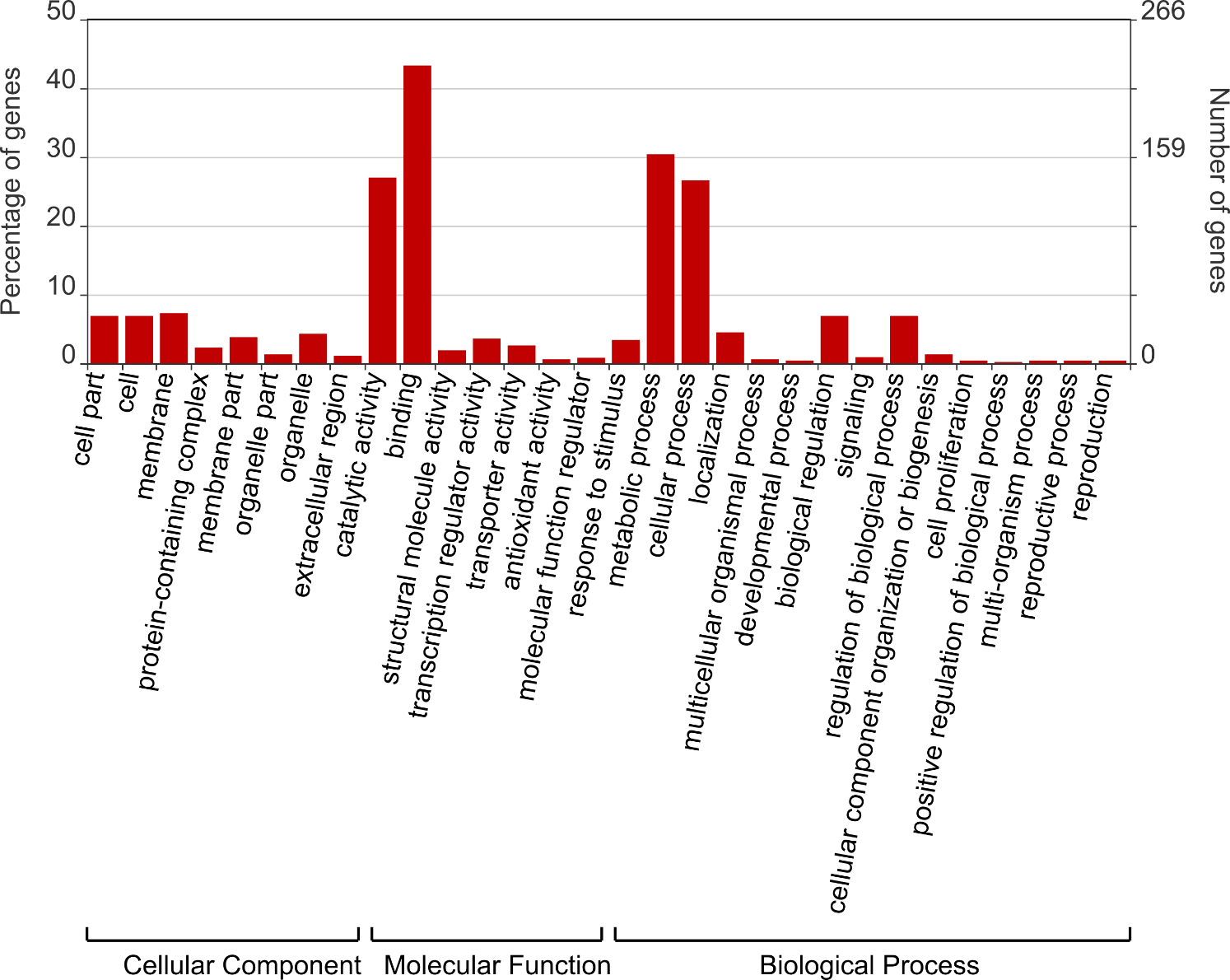
**Supplementary Figure S6.** Logarithm of the -values [] profiles from random-effect multiple interval mapping (REMIM) for seven B2721 traits evaluated for four years (2006-8 and 2014). QTL peak locations are represented by triangles and their ~95% support intervals by light-shaded rectangles. Traits: plant yield (PY), foliage maturity (FM), specific gravity (SG), dry matter (DM), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI). (PDF)



**Supplementary Figure S7.** Logarithm of the odds (LOD) profiles from fixed-effect interval mapping (FEIM) for seven B2721 traits evaluated for four years (2006-8 and 2014). QTL peak locations are represented by triangles and their ~95% support intervals by light-shaded rectangles. Dashed lines denote the permutation-based thresholds. Traits: plant yield (PY), foliage maturity (FM), specific gravity (SG), dry matter (DM), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI). (PDF)



**Supplementary Figure S8.** Correlations between adjusted means (upper diagonal) and QTL-based predicted means (lower diagonal) for B2721 traits (only those with identified QTL) evaluated for four years (2006-8 and 2014). Traits: plant yield (PY), foliage maturity (FM), specific gravity (SG), dry matter (DM), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI). (PDF)



**Supplementary Figure S9**. Enriched Gene Ontology (GO) terms for 533 annotated genes within our QTL regions. (PDF)

## Supplementary Tables

**Supplementary Table S1.** B2721 genetic map summary.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Linkage  group | Length  (cM) | Number of SNPs | | | Total | SNPs/cM |
| Simplex | Double-  simplex | Multiplex |
| 1 | 205.88 | 102 | 122 | 252 | 476 | 2.31 |
| 2 | 125.11 | 131 | 139 | 160 | 430 | 3.44 |
| 3 | 134.07 | 151 | 21 | 210 | 382 | 2.85 |
| 4 | 165.9 | 114 | 87 | 224 | 425 | 2.56 |
| 5 | 106.2 | 124 | 53 | 134 | 311 | 2.93 |
| 6 | 142.7 | 73 | 75 | 249 | 397 | 2.78 |
| 7 | 126.09 | 136 | 94 | 184 | 414 | 3.28 |
| 8 | 136.68 | 61 | 98 | 183 | 342 | 2.5 |
| 9 | 144.49 | 76 | 101 | 179 | 356 | 2.46 |
| 10 | 114.67 | 47 | 36 | 124 | 207 | 1.81 |
| 11 | 110.66 | 95 | 52 | 173 | 320 | 2.89 |
| 12 | 117.54 | 115 | 32 | 78 | 225 | 1.91 |
| Total | 1629.99 | 1225 | 910 | 2150 | 4285 | 2.64 |

**Supplementary Table S2.** IlluminaInfinium® 8,303 Potato Array marker names (solcap\_snp) and their respective *Solanum tuberosum* v. 4.03 (ST4.03) genome positions, in base pairs (bp), on the left and on the right of the QTL peak and their support intervals.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Traita | QTL | Linkage  group | QTL  peak  (cM) | Support  interval  (cM) | QTL peak | | | | Support interval | | | |
| Left marker | | Right marker | | Left marker | | Right marker | |
| solcap\_snp | ST4.03 | solcap\_snp | ST4.03 | solcap\_snp | ST4.03 | solcap\_snp | ST4.03 |
| PY06 | 1 | 5 | 32 | 13-42 | c2\_50305 | 5,052,466 | c1\_14801 | 5,053,838 | c2\_11737 | 2,068,427 | c1\_5836 | 8,809,052 |
| PY08 | 1 | 5 | 46 | 0-53 | c2\_43535 | 9,273,361 | c2\_55894 | 9,558,289 | c2\_23776 | 65,193 | c1\_15638 | 13,087,731 |
| FM07 | 1 | 5 | 26 | 18-29 | c2\_11829 | 4,041,510 | c2\_22986 | 4,279,075 | c2\_11685 | 2,288,291 | c2\_23055 | 4,936,332 |
|  | 2 | 7 | 66 | 52-70 | c1\_10020 | 45,145,834 | c2\_33495 | 45,145,932 | c1\_7515 | 41,721,163 | c2\_45188 | 46,763,232 |
| FM08 | 1 | 1 | 9 | 0-32 | c2\_6713 | 2,068,305 | c2\_21097 | 2,589,277 | c2\_51460 | 151,047 | c2\_27877 | 6,071,374 |
|  | 2 | 5 | 27 | 0-41 | c2\_22959 | 4,434,048 | c2\_23052 | 4,906,728 | c2\_23776 | 65,193 | c1\_5836 | 8,809,052 |
|  | 3 | 7 | 65 | 35-71 | c2\_44105 | 44,499,261 | c2\_44120 | 44,687,759 | c2\_52374 | 5,791,166 | c2\_45182 | 46,813,301 |
| FM14 | 1 | 5 | 18 | 0-29 | c2\_11685 | 2,288,291 | c1\_3840 | 3,134,967 | c2\_23776 | 65,193 | c2\_23055 | 4,936,332 |
|  | 2 | 7 | 66 | 54-70 | c1\_10020 | 45,145,834 | c2\_33495 | 45,145,932 | c2\_23347 | 42,120,645 | c2\_45188 | 46,763,232 |
| SG06 | 1 | 2 | 97 | 34-117 | c2\_22890 | 42,663,931 | c2\_22853 | 42,733,475 | c1\_11494 | 26,262,335 | c2\_24869 | 46,387,242 |
|  | 2 | 3 | 44 | 31-50 | c2\_36469 | 38,175,099 | c1\_10879 | 38,177,443 | c1\_16267 | 17,239,217 | c1\_10514 | 40,768,867 |
| SG07 | 1 | 5 | 19 | 4-32 | c1\_3840 | 3,134,967 | c1\_3803 | 3,585,641 | c2\_23846 | 727,316 | c1\_14801 | 5,053,838 |
| SG08 | 1 | 5 | 28 | 15-32 | c2\_22959 | 4,434,048 | c2\_23052 | 4,906,728 | c2\_11712 | 2,149,871 | c1\_14801 | 5,053,838 |
| ST07 | 1 | 9 | 41 | 27-82 | c2\_3962 | 4,889,182 | c2\_52241 | 5,790,649 | c2\_39216 | 2,792,921 | c2\_19556 | 48,069,790 |
| ST08 | 1 | 4 | 2 | 0-4 | c1\_7574 | 117,620 | c2\_23593 | 635,886 | c1\_7574 | 117,620 | c2\_23600 | 656,192 |
| NS06 | 1 | 1 | 0 | 0-52 | c2\_51460 | 151,047 | c2\_51460 | 151,047 | c2\_51460 | 151,047 | c2\_55008 | 12,994,277 |
| NS08 | 1 | 5 | 26 | 12-32 | c2\_11829 | 4,041,510 | c2\_22986 | 4,279,075 | c2\_52084 | 1,902,218 | c1\_14801 | 5,053,838 |
| NI06 | 1 | 1 | 0 | 0-69 | c2\_51460 | 151,047 | c2\_51460 | 151,047 | c2\_51460 | 151,047 | c2\_2721 | 58,197,448 |
| NI07 | 1 | 1 | 0 | 0-8 | c2\_51460 | 151,047 | c2\_51460 | 151,047 | c2\_51460 | 151,047 | c2\_6906 | 1,606,674 |
|  | 2 | 5 | 25 | 8-34 | c2\_11829 | 4,041,510 | c2\_22986 | 4,279,075 | c2\_33543 | 1,505,540 | c2\_52436 | 6,147,926 |
| NI08 | 1 | 5 | 27 | 12-32 | c2\_22959 | 4,434,048 | c2\_23052 | 4,906,728 | c2\_52084 | 1,902,218 | c1\_14801 | 5,053,838 |

aTraits: plant yield (PY), foliage maturity (FM), specific gravity (SG), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI).

**Supplementary Table S3.** Fixed-effect interval mapping (FEIM) for B2721 traits evaluated for four years (2006-8 and 2014).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Traita | QTL | Linkage  group | Position  (cM) | Left marker | | Right marker | | LRTb | LODc | d  (%) |
| solcap\_snp | ST4.03 | solcap\_snp | ST4.03 |
| PY08 | 1 | 5 | 45 | c2\_43535 | 9,273,361 | c2\_55894 | 9,558,289 | 39.05 | 8.48 | 20.5 |
| FM07 | 1 | 5 | 26 | c2\_11829 | 4,041,510 | c2\_22986 | 4,279,075 | 93.00 | 20.19 | 46.0 |
|  | 2 | 7 | 66 | c1\_10020 | 45,145,834 | c2\_33495 | 45,145,932 | 35.01 | 7.60 | 18.5 |
| FM08 | 1 | 5 | 28 | c2\_22959 | 4,434,048 | c2\_23052 | 4,906,728 | 113.12 | 24.56 | 52.7 |
| FM14 | 1 | 5 | 26 | c2\_11829 | 4,041,510 | c2\_22986 | 4,279,075 | 56.34 | 12.24 | 30.1 |
|  | 2 | 7 | 66 | c1\_10020 | 45,145,834 | c2\_33495 | 45,145,932 | 29.63 | 6.43 | 15.4 |
| SG06 | 1 | 2 | 42 | c2\_41963 | 27,511,649 | c2\_41975 | 27,547,993 | 26.67 | 5.79 | 13.4 |
|  | 2 | 8 | 90 | c2\_7353 | 48,895,181 | c1\_13116 | 49,071,663 | 28.72 | 6.24 | 14.6 |
|  | 3 | 10 | 65 | c2\_27827 | 50,697,563 | c2\_27829 | 50,782,097 | 29.35 | 6.37 | 15.0 |
| SG07 | 1 | 5 | 21 | c1\_3840 | 3,134,967 | c1\_3803 | 3,585,641 | 31.48 | 6.83 | 16.4 |
| SG08 | 1 | 5 | 29 | c2\_23052 | 4,906,728 | c2\_23055 | 4,936,332 | 57.98 | 12.59 | 30.4 |
| ST08 | 1 | 4 | 2 | c1\_7574 | 117,620 | c2\_23593 | 635,886 | 56.21 | 12.21 | 29.5 |
|  | 2 | 9 | 92 | c2\_14640 | 50,367,678 | c2\_14641 | 50,367,972 | 27.13 | 5.89 | 13.6 |
| NS06 | 1 | 1 | 2 | c2\_36664 | 535,454 | c2\_36668 | 559,640 | 26.02 | 5.65 | 13.0 |
| NS08 | 1 | 5 | 27 | c2\_22959 | 4,434,048 | c2\_23052 | 4,906,728 | 26.45 | 5.74 | 13.2 |
| NI08 | 1 | 5 | 28 | c2\_22959 | 4,434,048 | c2\_23052 | 4,906,728 | 30.37 | 6.60 | 15.6 |
| NI14 | 1 | 5 | 27 | c2\_22959 | 4,434,048 | c2\_23052 | 4,906,728 | 26.42 | 5.74 | 13.5 |

aTraits: plant yield (PY), foliage maturity (FM), specific gravity (SG), skin texture (ST), and internal heat necrosis severity (NS) and intensity (NI).

bLikelihood ratio test (LRT).

cLogarithm of the odds (LOD).

dAdjusted R-squared () in percentage.

## Supplementary Files

**Supplementary File S1**. B2721 adjusted means for seven traits evaluated in four years (2006-8 and 2014). (XLXS)

**Supplementary File S2**. Correlation between B2721 mapping population traits based on pairwise complete observations. (XLXS)

**Supplementary File S3**. B2721 normalized intensities from Illumina Infinium® 8,303 Potato Array. (CSV)

**Supplementary File S4**. Illumina Infinium® 8,303 Potato Array sequence BLAST alignment against the *Solanum tuberosum* genome v. 4.03. (CSV)

**Supplementary File S5**. Linkage map information for B2721 mapping population. (XLXS)

**Supplementary File S6**. List of 533 annotated genes from *Solanum tuberosum* v. 4.03 genome related with B2721 QTL regions. (XLXS)

# References

Bethke, P. C., Nassar, A. M. K., Kubow, S., Leclerc, Y. N., Li, X. Q., Haroon, M., et al. (2014). History and Origin of Russet Burbank (Netted Gem) a Sport of Burbank. *Am. J. Potato Res.* 91, 594–609. doi:10.1007/s12230-014-9397-5.

Bourke, P. M., van Geest, G., Voorrips, R. E., Jansen, J., Kranenburg, T., Shahin, A., et al. (2018). polymapR—linkage analysis and genetic map construction from F1 populations of outcrossing polyploids. *Bioinformatics*, 1–7. doi:10.1093/bioinformatics/bty371.

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., and Maliepaard, C. (2015). The double-reduction landscape in tetraploid potato as revealed by a high-density linkage map. *Genetics* 201, 853–863. doi:10.1534/genetics.115.181008.

Bradshaw, J. E., Hackett, C. A., Pande, B., Waugh, R., and Bryan, G. J. (2008). QTL mapping of yield, agronomic and quality traits in tetraploid potato (Solanum tuberosum subsp. tuberosum). *Theor. Appl. Genet.* 116, 193–211. doi:10.1007/s00122-007-0659-1.

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., and Thompson, R. (2018). *ASReml-R Reference Manual Version 4*. Hemel Hempstead Available at: http://www.homepages.ed.ac.uk/iwhite/asreml/uop.

Chen, J., Zhang, F., Wang, L., Leach, L., and Luo, Z. (2018). Orthogonal contrast based models for quantitative genetic analysis in autotetraploid species. *New Phytol.* 220, 332–346. doi:10.1111/nph.15284.

Choudhary, A., Wright, L., Ponce, O., Chen, J., Prashar, A., Sanchez-Moran, E., et al. (2020). Varietal variation and chromosome behaviour during meiosis in Solanum tuberosum. *Heredity (Edinb).* doi:10.1038/s41437-020-0328-6.

Churchill, G. A., and Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 138, 963–71.

D’hoop, B. B., Keizer, P. L. C., Paulo, M. J., Visser, R. G. F., van Eeuwijk, F. A., and van Eck, H. J. (2014). Identification of agronomically important QTL in tetraploid potato cultivars using a marker-trait association analysis. *Theor. Appl. Genet.* 127, 731–748. doi:10.1007/s00122-013-2254-y.

da Silva Pereira, G., Gemenet, D. C., Mollinari, M., Olukolu, B. A., Wood, J. C., Diaz, F., et al. (2020). Multiple QTL Mapping in Autopolyploids: A Random-Effect Model Approach with Application in a Hexaploid Sweetpotato Full-Sib Population. *Genetics* 215, 579–595. doi:10.1534/genetics.120.303080.

Danan, S., Veyrieras, J. B., and Lefebvre, V. (2011). Construction of a potato consensus map and QTL meta-analysis offer new insights into the genetic architecture of late blight resistance and plant maturity traits. *BMC Plant Biol.* 11. doi:10.1186/1471-2229-11-16.

Dixit, S., Kumar Biswal, A., Min, A., Henry, A., Oane, R. H., Raorane, M. L., et al. (2015). Action of multiple intra-QTL genes concerted around a co-localized transcription factor underpins a large effect QTL. *Sci. Rep.* 5, 1–16. doi:10.1038/srep15183.

Dumble, S., Bernal, E. F., and Villardon, P. G. (2017). GGEBiplots: GGE Biplots with “ggplot2.” Available at: https://cran.r-project.org/package=GGEBiplots.

FAO (2020). FAOSTAT Crops. Available at: http://www.fao.org/faostat/ [Accessed January 23, 2020].

Felcher, K. J., Coombs, J. J., Massa, A. N., Hansey, C. N., Hamilton, J. P., Veilleux, R. E., et al. (2012). Integration of two diploid potato linkage maps with the potato genome sequence. *PLoS One* 7, e36347. doi:10.1371/journal.pone.0036347.

Ghislain, M., and Douches, D. S. (2020). “The Genes and Genomes of the Potato,” in *The Potato Crop*, eds. H. Campos and O. Ortiz (Springer), 139–162. doi:10.1007/978-3-030-28683-5\_5.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178--D1186. doi:10.1093/nar/gkr944.

Hackett, C. A., Boskamp, B., Vogogias, A., Preedy, K. F., Milne, I., Vogogias, T., et al. (2016). TetraploidSNPMap: software for linkage analysis and QTL mapping in autotetraploid populations using SNP dosage data. *J. Hered.* 108, 438–442. doi:10.1093/jhered/esx022.

Hackett, C. A., Bradshaw, J. E., and Bryan, G. J. (2014). QTL mapping in autotetraploids using SNP dosage information. *Theor. Appl. Genet.* doi:10.1007/s00122-014-2347-2.

Hämäläinen, J. H., Watanabe, K. N., Valkonen, J. P. T., Arihara, A., Plaisted, R. L., Pehu, E., et al. (1997). Mapping and marker-assisted selection for a gene for extreme resistance to potato virus Y. *Theor. Appl. Genet.* 94, 192–197. doi:10.1007/s001220050399.

He, X., and Zhang, J. (2006). Toward a molecular understanding of pleiotropy. *Genetics* 173, 1885–1891. doi:10.1534/genetics.106.060269.

Jong, H. (1981). Inheritance of russeting in cultivated diploid potatoes. *Potato Res.* 24, 309–313.

Kempthorne, O. (1955). The correlation between relatives in a simple autotetraploid population. *Genetics* 40, 168–174.

Kloosterman, B., Abelenda, J. A., Gomez, M. D. M. C., Oortwijn, M., De Boer, J. M., Kowitwanich, K., et al. (2013). Naturally occurring allele diversity allows potato cultivation in northern latitudes. *Nature* 495, 246–250. doi:10.1038/nature11912.

Li, W., Dong, J., Cao, M., Gao, X., Wang, D., Liu, B., et al. (2019). Genome-wide identification and characterization of HD-ZIP genes in potato. *Gene* 697, 103–117. doi:10.1016/j.gene.2019.02.024.

Luo, Z. W., Zhang, R. M., and Kearsey, M. J. (2004). Theoretical basis for genetic linkage analysis in autotetraploid species. *PNAS* 101, 7040–7045.

Mann, H., Iorizzo, M., Gao, L., D’Agostino, N., Carputo, D., Chiusano, M. L., et al. (2011). “Molecular Linkage Maps: Strategies, Resources and Achievements,” in *Genetics, Genomics and Breeding of Potato*, eds. J. M. Bradeen and C. Kole (Enfield: CRC Press), 68–89.

Manrique-Carpintero, N. C., Coombs, J. J., Cui, Y., Veilleux, R. E., Robin Buell, C., and Douches, D. (2015). Genetic map and QTL analysis of agronomic traits in a diploid potato population using single nucleotide polymorphism markers. *Crop Sci.* 55, 2566–2579. doi:10.2135/cropsci2014.10.0745.

Massa, A. N., Manrique-Carpintero, N. C., Coombs, J., Haynes, K. G., Bethke, P. C., Brandt, T. L., et al. (2018). Linkage analysis and QTL mapping in a tetraploid russet mapping population of potato. *BMC Genet.* 19, 1–13. doi:10.1186/s12863-018-0672-1.

McCord, P. H., Sosinski, B. R., Haynes, K. G., Clough, M. E., and Yencho, G. C. (2011a). Linkage mapping and QTL analysis of agronomic traits in tetraploid potato (Solanum tuberosum subsp. tuberosum). *Crop Sci.* 51, 771–785. doi:10.2135/cropsci2010.02.0108.

McCord, P. H., Sosinski, B. R., Haynes, K. G., Clough, M. E., and Yencho, G. C. (2011b). QTL mapping of internal heat necrosis in tetraploid potato. *Theor. Appl. Genet.* 122, 129–142. doi:10.1007/s00122-010-1429-z.

Mizoi, J., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1819, 86–96. doi:10.1016/j.bbagrm.2011.08.004.

Mollinari, M., and Garcia, A. A. F. (2019). Linkage Analysis and Haplotype Phasing in Experimental Autopolyploid Populations with High Ploidy Level Using Hidden Markov Models. *G3 Genes|Genomes|Genetics* 9, 3297–3314. doi:10.1534/g3.119.400378.

Mollinari, M., Olukolu, B. A., Pereira, G. S., Khan, A., Gemenet, D., Yencho, G. C., et al. (2020). Unraveling the Hexaploid Sweetpotato Inheritance Using Ultra-Dense Multilocus Mapping. *G3 Genes|Genomes|Genetics* 10, 281–292. doi:10.1534/g3.119.400620.

Olsen, A. N., Ernst, H. A., Leggio, L. Lo, and Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci.* 10, 79–87. doi:10.1016/j.tplants.2004.12.010.

Preedy, K. F., and Hackett, C. A. (2016). A rapid marker ordering approach for high-density genetic linkage maps in experimental autotetraploid populations using multidimensional scaling. *Theor. Appl. Genet.* 129, 2117–2132. doi:10.1007/s00122-016-2761-8.

Qu, L., Guennel, T., and Marshall, S. L. (2013). Linear score tests for variance components in linear mixed models and applications to genetic association studies. *Biometrics* 69, 883–892. doi:10.1111/biom.12095.

R Core Team (2019). R: A Language and Environment for Statistical Computing. Available at: http://www.r-project.org.

Rak, K., Bethke, P. C., and Palta, J. P. (2017). QTL mapping of potato chip color and tuber traits within an autotetraploid family. *Mol. Breed.* 37. doi:10.1007/s11032-017-0619-7.

Rosin, F. M., Hart, J. K., Horner, H. T., Davies, P. J., and Hannapel, D. J. (2003). Overexpression of a Knotted-Like Homeobox Gene of Gibberellin Accumulation 1. *Plant Physiol.* 132, 106–117. doi:10.1104/pp.102.015560.ser.

Ruiz, E., Jackson, S., and Cimentada, J. (2019). corrr: Correlations in R. Available at: https://cran.r-project.org/package=corrr.

Schmitz Carley, C. A., Coombs, J. J., Douches, D. S., Bethke, P. C., Palta, J. P., Novy, R. G., et al. (2017). Automated tetraploid genotype calling by hierarchical clustering. *Theor. Appl. Genet.* 130, 717–726. doi:10.1007/s00122-016-2845-5.

Schultz, L., Cogan, N. O. I., Mclean, K., Dale, M. F. B., Bryan, G. J., Forster, J. W., et al. (2012). Evaluation and implementation of a potential diagnostic molecular marker for H1-conferred potato cyst nematode resistance in potato (Solanum tuberosum L.). *Plant Breed.* 131, 315–321. doi:10.1111/j.1439-0523.2012.01949.x.

Schumann, M. J., Zeng, Z. B., Clough, M. E., and Yencho, G. C. (2017). Linkage map construction and QTL analysis for internal heat necrosis in autotetraploid potato. *Theor. Appl. Genet.* 130, 2045–2056. doi:10.1007/s00122-017-2941-1.

Shaner, G., and Finney, R. E. (1977). The Effect of Nitrogen Fertilization on the Expression of Slow-Mildewing Resistance in Knox Wheat. *Phytopathology* 67, 1051–1056. doi:10.1094/phyto-67-1051.

Sharma, S. K., Bolser, D., de Boer, J., Sonderkaer, M., Amoros, W., Carboni, M. F., et al. (2013). Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. *G3 Genes|Genomes|Genetics* 3, 2031–47. doi:10.1534/g3.113.007153.

Singh, A. K., Sharma, V., Pal, A. K., Acharya, V., and Ahuja, P. S. (2013). Genome-wide organization and expression profiling of the NAC transcription factor family in potato (solanum tuberosum L.). *DNA Res.* 20, 403–423. doi:10.1093/dnares/dst019.

Slater, A. T., Cogan, N. O. I., Hayes, B. J., Schultz, L., Dale, M. F. B., Bryan, G. J., et al. (2014). Improving breeding efficiency in potato using molecular and quantitative genetics. *Theor. Appl. Genet.* 127, 2279–2292. doi:10.1007/s00122-014-2386-8.

Soler, M., Verdaguer, R., Fernández-Piñán, S., Company-Arumí, D., Boher, P., Góngora-Castillo, E., et al. (2020). Silencing against the conserved NAC domain of the potato StNAC103 reveals new NAC candidates to repress the suberin associated waxes in phellem. *Plant Sci.* 291, 110360. doi:10.1016/j.plantsci.2019.110360.

Webb, R. E., Wilson, D. R., Shumaker, J. R., Graves, B., Henninger, M. R., Watts, J., et al. (1978). Atlantic: A new potato variety with high solids, good processing quality, and resistance to pests. *Am. Potato J.* 55, 141–145. doi:10.1007/BF02852087.

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., and Zemla, J. (2017). corrplot: Visualization of a Correlation Matrix. Available at: https://cran.r-project.org/package=corrplot.

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., et al. (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475, 189–195. doi:10.1038/nature10158.

Yan, W. H., Wang, P., Chen, H. X., Zhou, H. J., Li, Q. P., Wang, C. R., et al. (2011). A major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol. Plant* 4, 319–330. doi:10.1093/mp/ssq070.

Yan, W., and Kang, M. S. (2003). *GGE Biplot Analysis: a graphical tool for breeders, geneticists, and agronomists*. Boca Raton: CRC Press doi:10.1201/9781420040371.

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., et al. (2018). WEGO 2.0: A web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Res.* 46, W71–W75. doi:10.1093/nar/gky400.

Yencho, G. C., McCord, P. H., Haynes, K. G., and Sterrett, S. B. R. (2008). Internal heat necrosis of potato - A review. *Am. J. Potato Res.* 85, 69–76. doi:10.1007/s12230-008-9008-4.

Zou, F., Fine, J. P., Hu, J., and Lin, D. Y. (2004). An efficient resampling method for assessing genome-wide statistical significance in mapping quantitative trait loci. *Genetics* 168, 2307–2316. doi:10.1534/genetics.104.031427.

Zych, K., Gort, G., Maliepaard, C. A., Jansen, R. C., and Voorrips, R. E. (2019). FitTetra 2.0 - Improved genotype calling for tetraploids with multiple population and parental data support. *BMC Bioinformatics* 20, 1–8. doi:10.1186/s12859-019-2703-y.

# Conflict of Interest Statement

*The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest*.