Full Title: Quantitative trait loci (QTL) underlying phenotypic variation in bioethanol-related processes in *Neurospora crassa*.

Short Title: Bioethanol QTL

Joshua C. Waters[1,2], Deval Jhaveri[2], Justin C. Biffinger[3], and Kwangwon Lee[1,2,*]

[1] Department of Biology, Rutgers, The State University of New Jersey - Camden, 315 Penn Street, Camden, NJ, United States of America

[2] Center for Computational & Integrative Biology, Rutgers, The State University of New Jersey - Camden, 200 Federal Street, Camden, NJ, United States of America

[3] Chemistry Department, University of Dayton, Dayton OH., 300 College Park, Dayton OH, United States of America

[*] Corresponding author

E-mail: kwagnwon.lee@rutgers.edu

Keywords: Fungal breeding, bioethanol, QTL, natural variation

Abstract

Bioethanol production from lignocellulosic biomass has received increasing attention over the past decade. Many attempts have been made to reduce the cost of bioethanol production by combining the separate steps of the process into a single-step process known as consolidated bioprocessing. This requires identification of organisms that can efficiently decompose lignocellulose to simple sugars and ferment the pentose and hexose sugars liberated to ethanol. There have been many attempts in engineering laboratory strains by adding new genes or modifying genes to expand the capacity of an industrial microorganism. There has been less attention in improving bioethanol-related processes utilizing natural variation existing in the natural ecotypes. In this study, we sought to identify genomic loci contributing to variation in saccharification of cellulose and fermentation of glucose in the fermenting cellulolytic fungus *Neurospora crassa* through quantitative trait loci (QTL) analysis. We identified one major QTL contributing to fermentation of glucose and multiple putative QTL's underlying saccharification. Understanding the natural variation of the major QTL gene would provide new insights in developing industrial microbes for bioethanol production.

Introduction

Fermentation of cellulosic biomass by microorganisms is a complex process requiring coordinated regulation of multiple pathways, including production and secretion of at least three distinct classes of catabolic enzymes (cellulase, hemicellulase, and lignin peroxidase enzymes), transport of substrates into the cell, followed by catabolism and fermentation of the resultant monomeric C5 and C6 sugars. Identification of organisms efficient in all requisite processes has been a major limitation toward realizing the vision of the consolidated bioprocessing (CBP) of biomass to ethanol for fuels [1–4]. While many researchers attempt to engineer existing strains by introducing the genes for pathways lacking in the given organisms, success has been limited due to the increased physiological burdens imposed by genetic manipulation. There is also a growing concern that genetic manipulations that are effective in laboratory strains may not be translatable to more robust strains used in industry [3, 5]. There are, however, organisms that already possess the genetic elements needed to perform all of the requisite processes, such as the model filamentous fungus *Neurospora crassa* [9].

These types of native microorganisms present excellent potential sources of useful industrial enzymes and products. Breeding has proven to be a reliable method for manipulating genomes to select for desired traits, as evidenced by the successes of plant breeding [6]. Selective mating has long been used in plant breeding to improve traits ranging from fruit size and yield to appearance or flavor [7, 8]. Recently, selective breeding has also been used in fungi in attempts to select for increased saccharification and fermentation potential [9].

59 The advent of computational tools such as genome-wide association studies
60 (GWAS) and quantitative trait loci (QTL) analysis have had drastic impacts on how plant
61 geneticists design new lineages with improved traits and qualities [10]. GWAS and QTL
62 analysis are computational tools for investigating the genetic elements underlying
63 quantitative traits. Unlike qualitative traits such as color, which are inherited discretely
64 and exhibit discontinuous variation, quantitative traits exhibit continuous variation and
65 cannot be discretely grouped [10]. GWAS and QTL use polymorphic genetic markers to
66 predict the extent to which a given genetic locus contributes to observed variation in
67 quantitative traits, such as plant height, and researchers have been able to use genetic
68 methods to identify combinations of desired alleles in individuals that are then used for
69 breeding, known as marker assisted selection (MAS).

70 QTL analysis and MAS have revolutionized plant genetics and should be equally
71 useful for improvement of quantitative traits in fungi, such as saccharification and
72 fermentation of cellulose. While a wealth of information is available on the effects of
73 gene deletions or overexpression of cellulolytic enzymes or fermentation pathways, there
74 have been no studies to characterize the variation resulting from allelic variation in
75 populations [11–18, 18, 19]. While gene deletions and overexpression studies can
76 implicate genetic elements in these processes, these manipulations increase cellular
77 burden due to unknown pleiotropic effects, and multiple manipulations can result in
78 compounding deficits [3, 5, 20].

79 The inherent complexity of lignocellulose metabolism allows for abundant
80 sources of variation that will affect CBP performance. Understanding how allelic
81 variation attenuates performance will lead to a more robust understanding of cellulose

82 metabolism and identify superior combinations of alleles for trait enhancement for CBP.
83 To this end, we chose to perform QTL analysis on a laboratory generated population of
84 the model filamentous fungus *Neurospora crassa*. 111 offspring and two parental strains
85 were genotyped by sequencing, evaluated for their ability to decompose cellulose and
86 ferment glucose, and subjected to QTL analysis. A major QTL was identified for
87 fermentation and multiple putative QTL's were identified for saccharification of
88 cellulose.
89

90 Materials and methods

91 Strains

92 Parental strains (FGSC2223 and FGSC4825) from the Fungal Genetics Stock Center
93 (FGSC) [26], and their cross progeny (designated as the N6 population) were grown from
94 long-term stock prior to each experiment on Vogel's minimal media slants (1X Vogel's
95 Salts, 2% Sucrose, 1.5% Agar, pH 5.8). Spore suspensions in High Glucose Liquid Media
96 (HGLM) (1X Vogel's salts, 2% glucose, 0.5% L-Arginine, pH 5.8) were used to generate
97 mycelial mats in petri plates, which were used to generate replicate mycelial pads for
98 each experiment using a bore punch.
99

100 Enzyme activity assay

101 A modified FPA assay was performed using 96-well plates as described by Camassola
102 and Dillon, in which secreted protein extracts were taken from wells in 6-well plates
103 containing 1% CMC broth (1X Vogel's salts, 1% CMC) 4 days after inoculation with 10

104 gauge mycelial pads [27]. Culture broth containing secreted enzyme was filtered and
105 centrifuged at 13.2k rpm to remove any fungal cells or debris. 50µL of supernatant was
106 added to 100µL of 50mM sodium acetate buffer pH 5.6 in a 96-well deep-well plate,
107 which was then equilibrated to 50°C for 5 min in a hot-water bath. A 5mg strip of
108 Whatman Grade No. 1 filter paper was submerged in the solution, and the plate was
109 incubated at 50°C for 60 min. After 60 min, 300µL of 3,5-dinitrosalicylic acid (DNS)
110 Reagent was added to stop the enzymatic reaction and visualize glucose equivalents. The
111 plate was incubated at 100°C for 10 min to develop color, then transferred to an ice bath
112 to stop color development. 100µL of Enzyme/DNS mixture was transferred to a clear-
113 bottom assay plate, diluted with 200µL of deionized H_2O, and absorbance was measured
114 at 545 nm. The concentration of reducing equivalents released was determined with
115 standard curves generated with glucose standards of known concentration. All samples
116 were performed in biological triplicates. The activity could not be quantified in
117 conventional Filter Paper Units because of the low concentration of secreted enzymes.
118 Therefore, these results were presented in the form of the concentration of reducing
119 equivalents released in each aliquot.
120

121 Fermentations and ethanol quantification

122 To characterize fermentation of glucose among the lab generated N6 first filial (F1)
123 generation and the parental strains (FGSC 2223 and FGSC4825), fermentation was
124 carried out in a 96-well format in deep-well plates. Replicates for each strain were
125 collected from mycelial mats grown from spore suspension in high glucose liquid media
126 (HGLM) with a 6 gauge punch and inoculated into 750µl of HGLM (2% glucose), sealed

127 with aluminum ThermowellTM seals, and allowed to ferment for 7 days in 12:12
128 Light/Dark conditions at 25°C. All samples were performed in biological quadruplicate.
129 After fermentation, 600 µL of media was recovered and cell debris was removed by
130 sequential centrifugation at 13.2k rpm for 5 min. The recovered supernatant was
131 aliquoted into HPLC vials for ethanol analysis by HPLC. HPLC quantitation was
132 performed using a Varian ProStar HPLC with a Varian ProStar Autosampler and a
133 Varian 356-LC Refractive Index Detector. An isocratic elution was used with an Agilent
134 Hi-Plex H^+ (300 mm x 7.7 mm ion-exchange column) with 5mM H_2SO_4 at a flow rate of
135 0.7 ml/min at 60°C. The concentration of ethanol present was determined from a standard
136 curve based on ethanol standards with a known concentration.
137

138 Qualitative Trait Loci (QTL) Analysis

139 To investigate the underlying sources of variation in saccharification and fermentation, a
140 qualitative trait loci (QTL) analysis was performed using R to identify candidate genes
141 contributing to the observed variation. 113 strains from a previously genotyped
142 population (N6 population) were screened for saccharification and fermentation as
143 previously described and single nucleotide polymorphisms (SNP's) were used to generate
144 a linkage map for QTL Analysis. Briefly, 50,000 SNP markers were filtered to 146
145 evenly distributed informative markers using Python and later Excel. First, all markers
146 that were not polymorphic between parents were removed, followed by any markers in
147 which one of the two parents was missing data. The resulting 4900 markers were
148 formatted for Excel for further filtering. Chi-square tests were performed for each marker
149 and those markers with unequal segregation among progeny (>20% disparity) were

150 removed, followed by markers with >10% missing data among progeny. The filtered
151 genotype data was combined with phenotype data from FPA and glucose fermentation
152 assays and formatted for R-QTL. The formatted data was imported into R-QTL as and
153 the create map function was used to generate a linkage map from a physical map based
154 on recombination frequencies. Markers were then hand curated to generate a linkage map
155 for all 7 chromosomes (Linkage groups) of *Neurospora crassa* with approximately 22
156 evenly spaced markers per chromosome.

157 After generating a linkage map for the N6 population, QTL scans were performed
158 for each phenotype. Composite interval mapping (CIM) was used to scan for single
159 QTL's using the maximum likelihood (em) , Haley-Knott regression (hk), and multiple
160 imputation (imp) methods in the R-QTL package, allowing for up to 4 covariate markers
161 within a 20-marker window. Likelihood odds ratio (LOD) thresholds were generated for
162 95% and 90% confidence intervals for each method based on 1000 permutations of the
163 data. Putative QTL were designated as major QTL if they explained greater than 10% of
164 the observed variation [28]. Peak marker positions for each chromosome (based on LOD
165 score) were used to identify linkage blocks between flanking markers to identify
166 chromosomal regions containing the suspected QTL, according to their physical
167 positions. These chromosomal regions were searched in the Fungi Database
168 (FungiDB.org) to identify all genes within the region to find candidate genes contributing
169 to the observed variation.

170

171 # Results

172 To test if there exist a natural variation of the bioethanol-related processes in one
173 mapping population N6, we performed a QTL analyses in N6. Significant variation was
174 observed among the N6 population for both saccharification of cellulose and
175 fermentation of glucose (Fig 1). Both traits demonstrated transgressive segregation, with
176 21 or 58 of 111 offspring outperforming the more proficient parent and 66 or 33 of 111
177 underperforming the lesser parent in the FPA assay or fermentation of glucose,
178 respectively.
179
180 **Fig 1. Natural variation in cellulolytic activity and fermentation among the N6**
181 **population.** Significant variation was observed for saccharification of cellulose (left axis)
182 and fermentation of glucose (right axis) among the N6 population (ANOVA $p=3.99x10^{-33}$
183 and $p=8.75x10^{-17}$ respectively). Red bars represent mean and 95% confidence interval of
184 the mean.
185
186 Due to the low level of recombination from only 1 meiotic event in a single cross,
187 generating a map with evenly spaced markers required filtering a large number of
188 polymorphic markers, reducing the power of QTL analysis. A linkage map based on
189 recombination frequencies was generated from a physical map of SNP markers with
190 approximately 22 evenly spaced markers on each of *N. crassa's* 7 chromosomes (Fig 2).
191
192 **Fig 2. Linkage map of SNP markers for the N6 population.** A linkage map created in
193 R-QTL from physical marker positions, with approximately 22 evenly distributed
194 markers per chromosome. Locations and distance are measured in centimorgans (cM).

195

196 A major QTL region contributing to variation in fermentation was identified on
197 Linkage Group (LG) I (361 Genes - 160 Annotated, 142 Hypothetical, 59 Unspecified)
198 within a 90% confidence interval (CI) (p-value=0.075) (Fig 3). Additionally, minor
199 QTL's were observed for fermentation on LG IV (344 Genes - 158 Annotated, 134
200 Hypothetical, 52 Unspecified), LG VI (62 Genes - 28 Annotated, 32 Hypothetical, 2
201 Unspecified), and LG VII (77 Genes - 30 Annotated, 37 Hypothetical, 10 Unspecified),
202 however, none of these were above the LOD threshold for 90% CI.

203

204 **Fig 3. QTL analysis of Fermentation of Glucose to Ethanol by the N6 population.**
205 Composite Interval Mapping single-QTL scan for fermentation of glucose using linear
206 regression (EM Algorithm). The x-axis is a measure of the logarithm of the odds ratio
207 (LOD). Blue line represents 90% confidence interval.

208

209 Using the fitqtl function, it was observed that the 4 QTL markers account for
210 ~26% of the observed variation in fermentation of ethanol. The QTL on LG I accounts
211 for 13% of the observed variation, while QTL on LG IV, VI and VII account for 4, 7, and
212 6% of the variation, respectively. Interestingly, marker segregation analysis revealed that
213 the parental genotypes contributing to greater ethanol production varied between
214 markers, with parental genotype A accounting for higher production at marker Nc-M878
215 on LG I and genotype B accounting for higher production at marker Nc-M4092 on LG VI
216 (Fig 4A and B).

217

218 **Fig 4. Phenotype Segregation at QTL Markers for Fermentation.** A) Phenotype
219 segregation at the QTL marker on LG I according to genotype. B) Phenotype segregation
220 at the QTL marker on LG VI according to genotype.
221
222 Of particular interest in the linkage block in LG I, indicated by QTL analysis for
223 fermentation, was a cluster of annotated glycolytic enzymes, sugar transporters, enzymes
224 involved in alternate fates of pyruvate, including lactate dehydrogenase enzymes, and
225 components of energy production machinery (Table 1). Alternatively, the only annotated
226 gene of interest from the QTL on LG VI was another sugar transporter, *sut-7*.
227
228 **Table 1. Candidate QTL genes for ethanol fermentation.**

Gene ID	Gene Name	Description	Linkage Group	Location
NCU00629	emp-3	6-phosphofructokinase	I	CM002236: 7,764,287 - 7,768,767 (-)
NCU00575	gki-2	glucokinase-2	I	CM002236: 7,953,683 - 7,958,506 (-)
NCU00742	glp-1	glycerol-3-phosphate dehydrogenase	I	CM002236: 7,360,766 - 7,362,918 (+)
NCU00720	tca-17	L-lactate dehydrogenase	I	CM002236: 7,427,034 - 7,428,944 (+)
NCU00775	tca-6	isocitrate dehydrogenase subunit 1	I	CM002236: 7,214,344 - 7,217,183 (-)
NCU00780	ldh-4	lactate dehydrogenase-4	I	CM002236: 7,202,837 - 7,204,284 (+)
NCU00606	null	short-chain alcohol dehydrogenase	I	CM002236: 7,858,582 - 7,859,522 (-)
NCU00801	cdt-1	MFS lactose permease, Sugar (and other) transporter	I	CM002236: 7,121,829 - 7,125,021 (-)
NCU00821	sut-15	sugar transporter-15	I	CM002236: 7,056,000 - 7,058,684 (-)

NCU00809	sut-26	sugar transporter-26	I	CM002236: 7,088,949 - 7,091,552 (+)
NCU05627	sut-7	sugar transporter-7	VI	CM002241: 1,275,879 - 1,278,130 (-)

229

230 Similarly, multiple QTL were observed for saccharification on LG II (184 Genes -
231 84 Annotated, 64 Hypothetical, 36 Unspecified), LG IV (184 Genes - 74 Annotated, 72
232 Hypothetical, 38 Unspecified), and LG VI (144 Genes – 44 Annotated, 54 Hypothetical,
233 16 Unspecified), although none were above the LOD threshold for 90% CI (Fig 5).
234 Combined, the putative QTL account for 19.5% of the total variation observed in
235 saccharification of cellulose. The major QTL on LG II accounts for 11% of the observed
236 variation, while minor QTL on LG VI and VI account for 4.6% and 5.7%, respectively.

237

238 **Fig 5. QTL analysis of Saccharification of Cellulose by the N6 population.** Composite
239 Interval Mapping single-QTL scan for FPA assay using linear regression (EM
240 Algorithm). The x-axis is a measure of the logarithm of the odds ratio (LOD).

241

242 Marker segregation analysis at QTL markers on LG II and LG IV revealed that
243 genotype B was the main contributor to increased potential for saccharification of
244 cellulose (Fig 6A and B). From this finding, we chose to investigate the potential for
245 epistatic interactions among the QTL markers for saccharification of cellulose using
246 R/QTL's Scantwo function. Scantwo analysis revealed that the QTL marker on LG II
247 (Nc-M1321) exhibited strong interactions with markers across all chromosomes, with the
248 strongest interactions with the adjacent marker on LG II and the putative QTL marker on
249 LG IV (Nc-M2984) (Fig 7A). Marker interaction plots demonstrated that the genotype B

250 specific increases in FPA activity only held for the interactions between QTL markers on
251 LG II and LG IV, but not with the interactions between adjacent markers on LG II (Fig
252 7B and C).
253
254 **Fig 6. Phenotype Segregation at QTL Markers for Saccharification.** A) Phenotype
255 segregation at the QTL marker on LG II according to genotype. B) Phenotype
256 segregation at the QTL marker on LG IV according to genotype.
257
258 **Fig 7. Saccharification QTL Marker Interactions.** (A) Scantwo QTL analysis for
259 epistatic marker interactions. Heatmap legend represents logarithm of the odds ratio
260 (LOD) scores for marker interactions. (B) Average phenotypes according to genotype at
261 sequential markers on LG II [Marker Nc-1321 is QTL marker on LG II]. (C) Average
262 phenotypes according to genotype at QTL markers on LG II and LG IV. Colors represent
263 genotype at marker indicated in legend, while letters on x-axis indicate genotype at
264 marker indicated on x-axis title. Error bars represent standard deviations.
265
266 Among those genes within the QTL regions for saccharification there were some
267 glucanase and xylanase enzymes, and transcription factors, which may have a role in
268 sensing or degrading cellulose. However, two major regulatory genes of interest were
269 also identified, *vib-1* and *xlr-1*, on LG II and LG IV respectively (Table 2).
270
271 **Table 2. Candidate QTL genes involved in saccharification of cellulose.**

Gene ID	Gene Name	Description	Linkage Group	Location
NCU03725	vib-1	VIB-1, variant 1	II	CM002240: 2,215,819 - 2,219,221 (+)
NCU06971	xlr-1	transcriptional activator xlnR	IV	CM002239: 4,706,112 - 4,709,560 (-)
NCU01050	gh61-4	endoglucanase II	II	CM002240: 2,769,232 - 2,770,004 (+)
NCU03641	gh3-1	Beta-glucosidase 2, variant 1	II	CM002240: 1,920,114 - 1,923,804 (+)
NCU01059	gh47-3	glycosyl hydrolase family 47 protein	II	CM002240: 2,790,429 - 2,792,645 (-)
NCU01050	gh61-4	endoglucanase II	II	CM002240: 2,769,232 - 2,770,004 (+)
NCU03657	gh63-1	mannosyl-oligosaccharide glucosidase, variant 1	II	CM002240: 1,992,478 - 1,995,662 (+)
NCU01080	gh64-1	glucanase B	II	CM002240: 2,848,665 - 2,850,629 (-)
NCU06961	gh28-2	Exopolygalacturonase	IV	CM002239: 4,669,933 - 4,671,737 (+)
NCU07326	gh43-6	hypothetical protein	IV	CM002239: 4,359,254 - 4,361,686 (-)
NCU07005	gh76-9	glycosyl hydrolase family 76-9	IV	CM002239: 4,865,028 - 4,867,282 (-)
NCU07269	gh92-2	alpha-1,2-mannosidase	IV	CM002239: 4,569,439 - 4,572,146 (-)

Discussion

In current study, we have identified that there exists a substantial variation in bioethanol-related traits in the mapping population, N6. We also identified few major candidate genes contributing to the observed variation, including the clustered glucose metabolism genes on LG I for fermentation, and transcriptional regulators *vib-1* and *xlr-1* on LG II

279 and LG IV, respectively, for saccharification. The clustering of glucose metabolism genes
280 on LG I makes identification of candidate genes for further analysis more complicated,
281 requiring gene expression analysis for identification of those genes within the region that
282 are differentially regulated and contributing to the observed variation in ethanol
283 fermentation potential. While it is possible that an increased rate of glycolysis is
284 responsible for increased fermentative capacity, the presence of lactate dehydrogenase
285 enzymes within the QTL regions could present an alternative answer, with decreased
286 lactic acid fermentation resulting in increased rates of ethanol fermentation.

287 The presence of major metabolic regulators within QTL regions for
288 saccharification provide more promising QTL candidate genes. The first major gene of
289 interest for saccharification of cellulose is the transcription factor, vegetative
290 incompatibility blocked-1 (*vib-1*), located on LG II, which is involved in sensing carbon
291 starvation and signaling for induction of the cellulolytic response [21]. The second major
292 gene of interest for saccharification of cellulose indicated is the xylose degradation
293 regulator (*xlr-1*) located on LG IV, as it is involved in induction of the hemicellulose
294 response, and regulation of cellodextrin transporters and cellulose degradation regulator
295 *clr-1*, along with *vib-1*, identified in the putative QTL on LG II [12, 22].

296 Although *xlr-1* has not been shown to directly bind to promoters for cellulase
297 enzymes, it has been implicated in cellulase production. In transcriptional analyses, two
298 independent groups demonstrated that *xlr-1* knockouts have reduced hemicellulase and
299 cellulase production [12, 17]. ChIP-seq experiments further revealed that although *xlr-1*
300 does not directly bind to cellulase promoters it does bind to the promoter of transcription
301 factor *clr-1*, a major regulator of cellulase expression, and can be found bound to targets

302 in conjunction with *clr-1* and/or *clr-2* [12]. Furthermore, *xlr-1* is involved in the
303 regulation of *vib-1*, which in turn further regulates expression of *clr-1* [12]. Similarly, Cai
304 et al., demonstrated that *xlr-1*, not *clr-1/2*, is required for expression of cellodextrin
305 transporter-2 (*cdt-2*), which is involved in sensing cellulose in the environment as well as
306 transmembrane transport [23]. *Cdt-2* has been implicated as a major cellodextrin importer
307 for facilitating induction of the cellulase response, as cellulase production and growth on
308 cellulose are significantly reduced in *cdt-2* knockouts [23, 24]. Overexpression of *cdt-2*
309 has also been shown to increase cellulase and hemicellulase production under cellulose
310 and xylan conditions [23]. These data suggest pleiotropic or epistatic roles for *xlr-1* as a
311 key component for cellulase production through its role in regulating of *vib-1* and *cdt-2*
312 for carbon sensing and cellulase induction, *clr-1* for inducing the cellulolytic response via
313 *clr-2*, and the co-regulation of genes targeted by *clr-1/2*.

314 Among other filamentous fungi, *xlr-1* orthologues (*xlnR/xyr-1*) have been shown
315 to directly regulate cellulase expression alongside hemicellulase, although this role has
316 been lost in *Neurospora* [12, 15, 17, 18, 25]. However, a vestige of this role may be
317 evident in *xlr-1*'s indirect role in the response to cellulose through its regulation of
318 critical genes involved in sensing cellulose in the environment, especially through *vib-1*,
319 *clr-1*, *cdt-2* and major facilitator superfamily (MFS) sugar transporters. While RNA-seq
320 and ChIP-seq experiments have previously been employed to decipher the role of *xlr-1* in
321 *N. crassa*, these studies have used knockout and overexpression strains that share the
322 same genetic scaffold as the wild type reference. Similar studies using diverse natural
323 strains may demonstrate that differential expression of *xlr-1* leads to differences in

cellulolytic potential or illustrate potentially disparate roles arising from allelic variation at the *xlr-1* locus or its targets.

Allelic variation in lignocellulose degradation regulators and their target genes may provide unique insights into the lignocellulose response of *N. crassa*. Transcriptomic profiling and expression analysis may reveal differential interactions between *xlr-1* and its' targets, resulting in differential enzyme induction and degradation of cellulosic substrates. Furthermore, expression QTL (eQTL) analysis would be important for discerning the pleiotropic effects of *xlr-1* from its epistatic effects. Understanding the implications of allelic effects on lignocellulose metabolism should allude to elite genotypes for further enhancement and possible industrial applications in bioethanol production.

Acknowledgements

The authors thank Drs. Sunil Shende and Charot Rodeget for curating Python scripts, critical reading and insightful comments.

Author Contributions

Conceptualization: Kwangwon Lee and Joshua C. Waters

Project Administration: Kwangwon Lee

Funding Acquisition: Kwangwon Lee

Methodology: Joshua C. Waters and Kwangwon Lee

Supervision: Joshua C. Waters and Kwangwon Lee

Investigation: Joshua C. Waters, Deval Jhaveri, Justin C. Biffinger

Validation: Joshua C. Waters, Deval Jhaveri, Justin C. Biffinger, Kwangwon Lee

Data curation: Deval Jhaveri

Software: Joshua C. Waters

Formal Analysis: Joshua C. Waters

Visualization: Joshua C. Waters

Writing-original draft: Joshua C. Waters and Kwangwon Lee

Writing-review and editing: Joshua C. Waters and Kwangwon Lee

Conflict of Interest Disclosure: The authors declare no conflict of interest.

References

1. Dashtban M, Schraft H, Qin W (2009) Fungal Bioconversion of Lignocellulosic Residues; Opportunities & Perspectives. *Int J Biol Sci* 5(6):578–595.

2. Dogaris I, Mamma D, Kekos D (2013) Biotechnological production of ethanol from renewable resources by Neurospora crassa: an alternative to conventional yeast fermentations? *Appl Microbiol Biotechnol* 97(4):1457–1473.

365 3. Favaro L, et al. (2013) Designing industrial yeasts for the consolidated bioprocessing of
366 starchy biomass to ethanol. *Bioengineered* 4(2):97–102.

367 4. Mattila H, Kuuskeri J, Lundell T (2017) Single-step, single-organism bioethanol production
368 and bioconversion of lignocellulose waste materials by phlebioid fungal species. *Bioresour*
369 *Technol* 225:254–261.

370 5. Zhang G-C, Turner TL, Jin Y-S (2017) Enhanced xylose fermentation by engineered yeast
371 expressing NADH oxidase through high cell density inoculums. *J Ind Microbiol Biotechnol*
372 44(3):387–395.

373 6. Hallauer AR, Carena MJ, Filho JBM (2010) *Quantitative Genetics in Maize Breeding*
374 (Springer-Verlag, New York). 3rd Ed. Available at:
375 https://www.springer.com/la/book/9781441907653 [Accessed June 27, 2019].

376 7. Muranty H, et al. (2015) Accuracy and responses of genomic selection on key traits in
377 apple breeding. *Hortic Res* 2:15060.

378 8. Qiu X, et al. (2017) Examining two sets of introgression lines reveals background-
379 independent and stably expressed QTL that improve grain appearance quality in rice
380 (Oryza sativa L.). *Theor Appl Genet*:1–17.

381 9. Waters JC, Nixon A, Dwyer M, Biffinger JC, Lee K (2017) Developing elite Neurospora crassa
382 strains for cellulosic ethanol production using fungal breeding. *J Ind Microbiol*
383 *Biotechnol*:1–8.

384 10. Kumar J, et al. (2017) Quantitative trait loci from identification to exploitation for crop
385 improvement. *Plant Cell Rep*:1–27.

386 11. Benz JP, et al. (2014) A comparative systems analysis of polysaccharide-elicited responses
387 in Neurospora crassa reveals carbon source-specific cellular adaptations. *Mol Microbiol*
388 91(2):275–299.

389 12. Craig JP, Coradetti ST, Starr TL, Glass NL (2015) Direct Target Network of the Neurospora
390 crassa Plant Cell Wall Deconstruction Regulators CLR-1, CLR-2, and XLR-1. *mBio*
391 6(5):e01452-15.

392 13. Hildebrand A, Szewczyk E, Lin H, Kasuga T, Fan Z (2015) Engineering Neurospora crassa for
393 Improved Cellobiose and Cellobionate Production. *Appl Environ Microbiol* 81(2):597–603.

394 14. Reilly MC, Qin L, Craig JP, Starr TL, Glass NL (2015) Deletion of homologs of the SREBP
395 pathway results in hyper-production of cellulases in Neurospora crassa and Trichoderma
396 reesei. *Biotechnol Biofuels* 8(1):1–18.

397 15. Stricker AR, Mach RL, Graaff LH de (2008) Regulation of transcription of cellulases- and
398 hemicellulases-encoding genes in Aspergillus niger and Hypocrea jecorina (Trichoderma
399 reesei). *Appl Microbiol Biotechnol* 78(2):211.

400 16. Sun J, Glass NL (2011) Identification of the CRE-1 Cellulolytic Regulon in Neurospora crassa.
401 *PLoS ONE* 6(9):e25654.

402 17. Sun J, Tian C, Diamond S, Glass NL (2012) Deciphering Transcriptional Regulatory
403 Mechanisms Associated with Hemicellulose Degradation in Neurospora crassa. *Eukaryot*
404 *Cell* 11(4):482–493.

405 18. Tani S, Kanamasa S, Sumitani J, Arai M, Kawaguchi T (2012) XlnR-independent signaling
406 pathway regulates both cellulase and xylanase genes in response to cellobiose in
407 Aspergillus aculeatus. *Curr Genet* 58(2):93–104.

408 19. Tian C, et al. (2009) Systems analysis of plant cell wall degradation by the model
409 filamentous fungus Neurospora crassa. *Proc Natl Acad Sci U S A* 106(52):22157–22162.

410 20. Sakuragi H, Kuroda K, Ueda M (2011) Molecular Breeding of Advanced Microorganisms for
411 Biofuel Production. *J Biomed Biotechnol* 2011. doi:10.1155/2011/416931.

412 21. Xiong Y, Sun J, Glass NL (2014) VIB1, a Link between Glucose Signaling and Carbon
413 Catabolite Repression, Is Essential for Plant Cell Wall Degradation by Neurospora crassa.
414 *PLOS Genet* 10(8):e1004500.

415 22. Benocci T, Aguilar-Pontes MV, Zhou M, Seiboth B, de Vries RP (2017) Regulators of plant
416 biomass degradation in ascomycetous fungi. *Biotechnol Biofuels* 10:152.

417 23. Cai P, et al. (2014) Evidence of a Critical Role for Cellodextrin Transporte 2 (CDT-2) in Both
418 Cellulose and Hemicellulose Degradation and Utilization in Neurospora crassa. *PLoS ONE*
419 9(2):e89330.

420 24. Galazka JM, et al. (2010) Cellodextrin Transport in Yeast for Improved Biofuel Production.
421 *Science* 330(6000):84–86.

422 25. Tani S, Kawaguchi T, Kobayashi T (2014) Complex regulation of hydrolytic enzyme genes
423 for cellulosic biomass degradation in filamentous fungi. *Appl Microbiol Biotechnol*
424 98(11):4829–4837.

26. McCluskey K, Wiest A, Plamann M (2010) The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. *J Biosci* 35(1):119–126.

27. Camassola M, J.P. Dillon A (2012) Cellulase Determination: Modifications to Make the Filter Paper Assay Easy, Fast, Practical and Efficient. *J Anal Bioanal Tech* 01(S1). doi:10.4172/scientificreports.125.

28. Hartman Y, Hooftman DAP, Eric Schranz M, van Tienderen PH (2013) QTL analysis reveals the genetic architecture of domestication traits in Crisphead lettuce. *Genet Resour Crop Evol* 60(4):1487–1500.

Supporting Information

S1 File. N6_QTL_Data. This .csv file contains genotype and phenotype information used for QTL analysis in R, using the R-qtl package.

S2 Text. N6 QTL script for analysis in R. This file contains R script used for QTL analysis in R, using the R-qtl package.

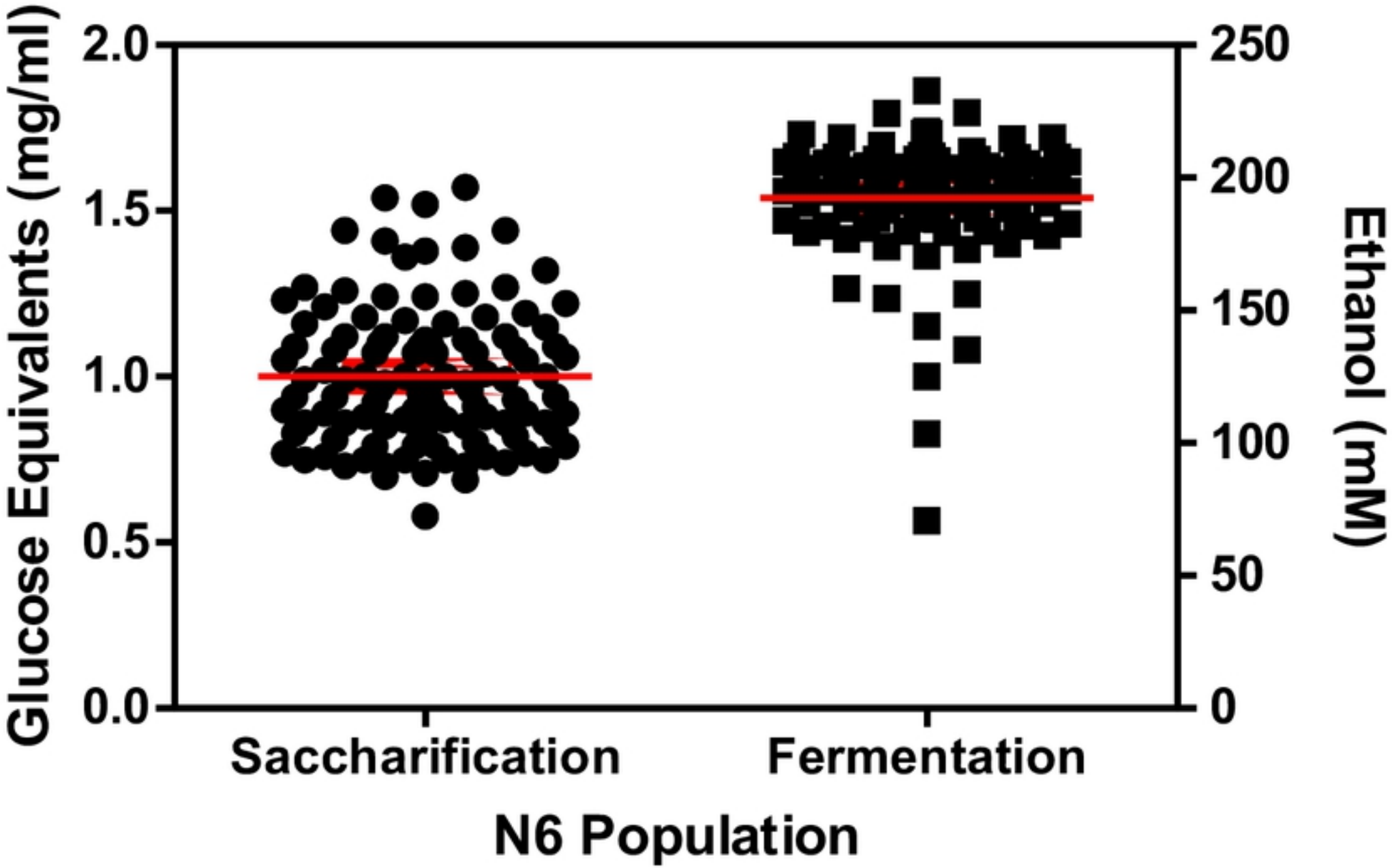
Glucose Equivalents (mg/ml)
2.0
1.5
1.0
0.5
0.0
250
200
150
100
50
0
Ethanol (mM)
Saccharification
Fermentation
N6 Population

Figure1

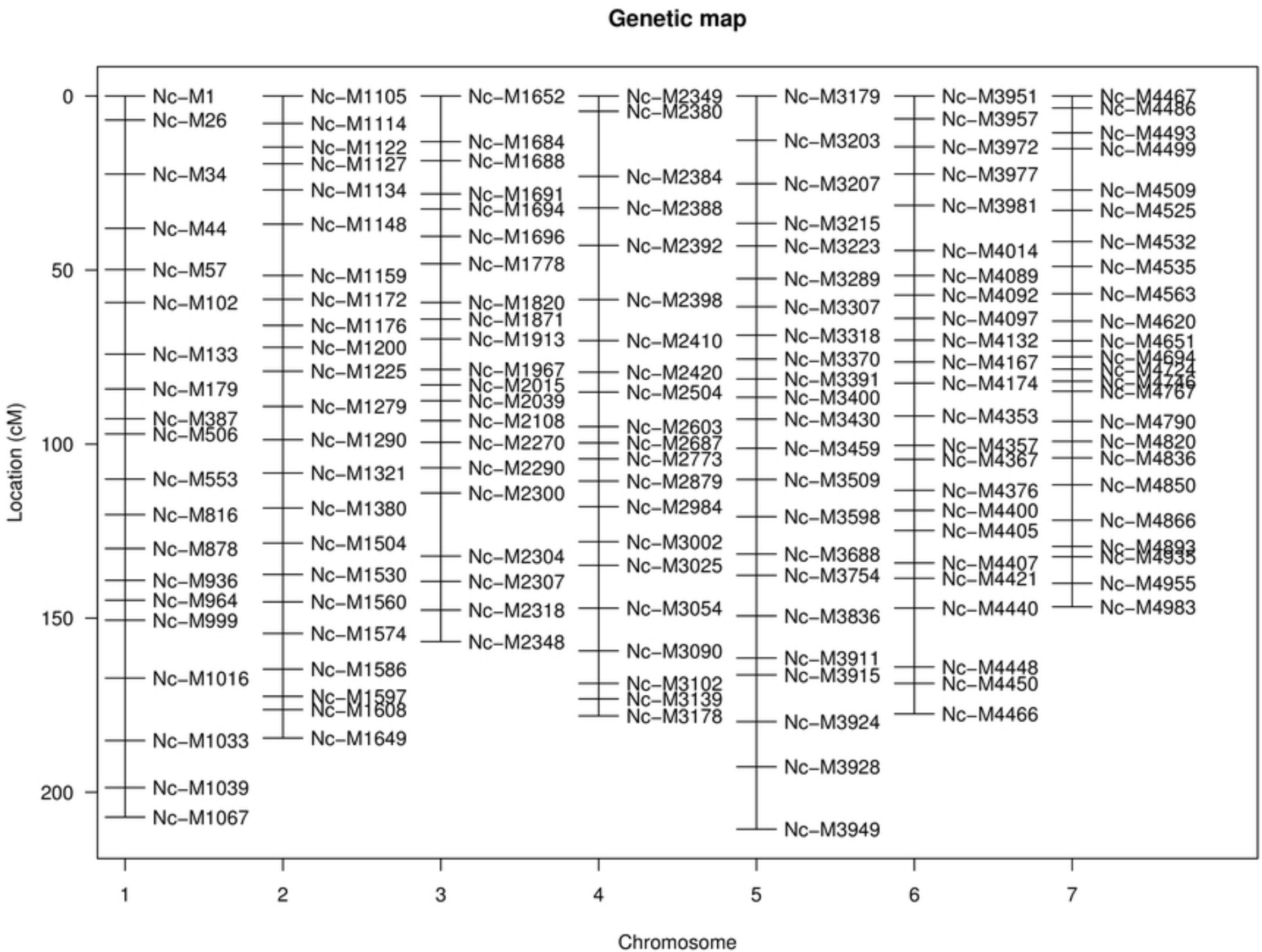
Genetic map
Location (cM)
0
50
100
150
200
Nc–M1
Nc–M26
Nc–M34
Nc–M44
Nc–M57
Nc–M102
Nc–M133
Nc–M179
Nc–M387
Nc–M506
Nc–M553
Nc–M816
Nc–M878
Nc–M936
Nc–M964
Nc–M999
Nc–M1016
Nc–M1033
Nc–M1039
Nc–M1067
Nc–M1105
Nc–M1114
Nc–M1122
Nc–M1127
Nc–M1134
Nc–M1148
Nc–M1159
Nc–M1172
Nc–M1176
Nc–M1200
Nc–M1225
Nc–M1279
Nc–M1290
Nc–M1321
Nc–M1380
Nc–M1504
Nc–M1530
Nc–M1560
Nc–M1574
Nc–M1586
Nc–M1597
Nc–M1608
Nc–M1649
Nc–M1652
Nc–M1684
Nc–M1688
Nc–M1691
Nc–M1694
Nc–M1696
Nc–M1778
Nc–M1820
Nc–M1871
Nc–M1913
Nc–M1967
Nc–M2015
Nc–M2039
Nc–M2108
Nc–M2270
Nc–M2290
Nc–M2300
Nc–M2304
Nc–M2307
Nc–M2318
Nc–M2348
Nc–M2349
Nc–M2380
Nc–M2384
Nc–M2388
Nc–M2392
Nc–M2398
Nc–M2410
Nc–M2420
Nc–M2504
Nc–M2603
Nc–M2687
Nc–M2773
Nc–M2879
Nc–M2984
Nc–M3002
Nc–M3025
Nc–M3054
Nc–M3090
Nc–M3102
Nc–M3139
Nc–M3178
Nc–M3179
Nc–M3203
Nc–M3207
Nc–M3215
Nc–M3223
Nc–M3289
Nc–M3307
Nc–M3318
Nc–M3370
Nc–M3391
Nc–M3400
Nc–M3430
Nc–M3459
Nc–M3509
Nc–M3598
Nc–M3688
Nc–M3754
Nc–M3836
Nc–M3911
Nc–M3915
Nc–M3924
Nc–M3928
Nc–M3949
Nc–M3951
Nc–M3957
Nc–M3972
Nc–M3977
Nc–M3981
Nc–M4014
Nc–M4089
Nc–M4092
Nc–M4097
Nc–M4132
Nc–M4167
Nc–M4174
Nc–M4353
Nc–M4357
Nc–M4367
Nc–M4376
Nc–M4400
Nc–M4405
Nc–M4407
Nc–M4421
Nc–M4440
Nc–M4448
Nc–M4450
Nc–M4466
Nc–M4467
Nc–M4486
Nc–M4493
Nc–M4499
Nc–M4509
Nc–M4525
Nc–M4532
Nc–M4535
Nc–M4563
Nc–M4620
Nc–M4651
Nc–M4694
Nc–M4724
Nc–M4746
Nc–M4767
Nc–M4790
Nc–M4820
Nc–M4836
Nc–M4850
Nc–M4866
Nc–M4893
Nc–M4935
Nc–M4955
Nc–M4983
1
2
3
4
5
6
7
Chromosome

Figure2

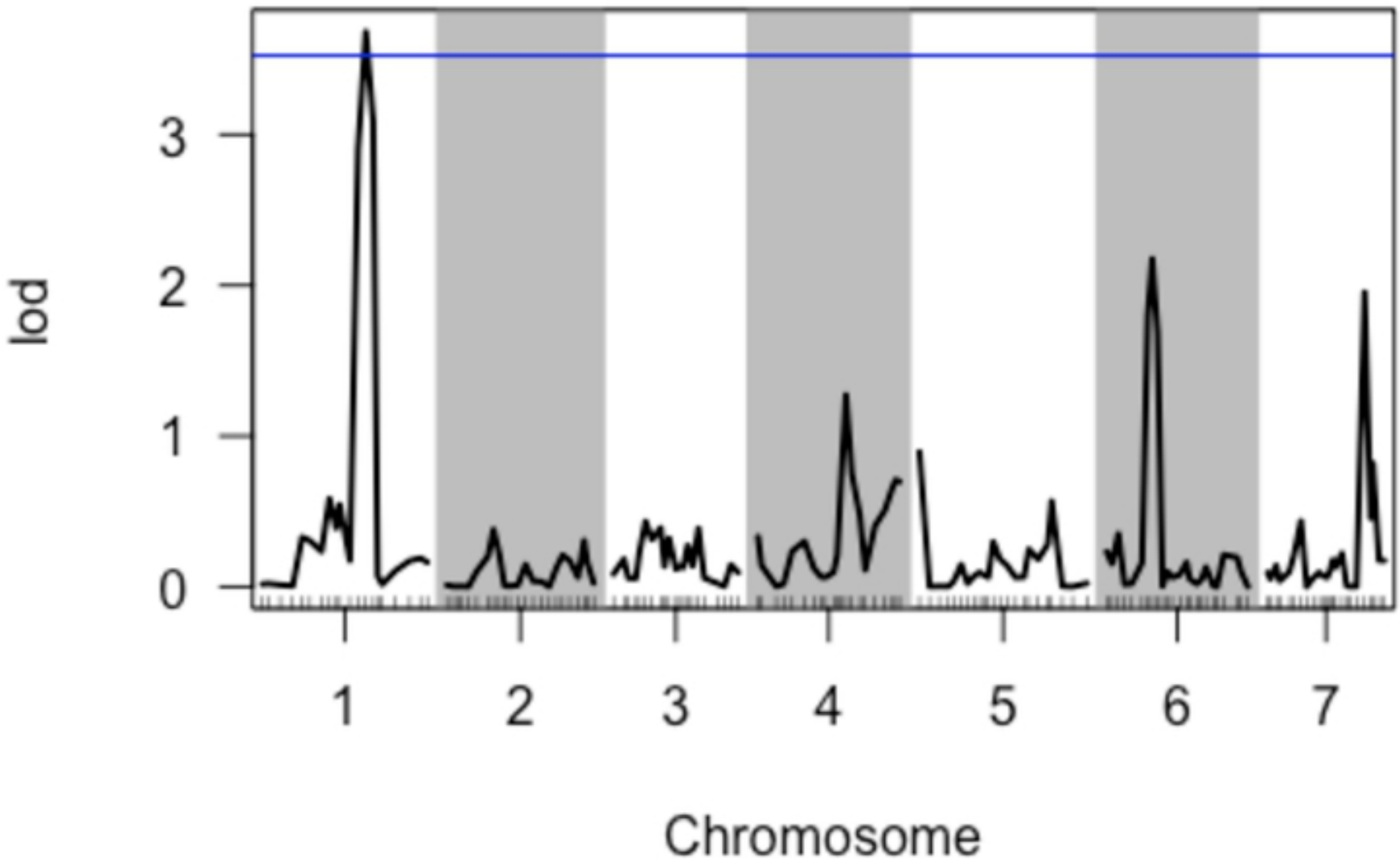

Figure3

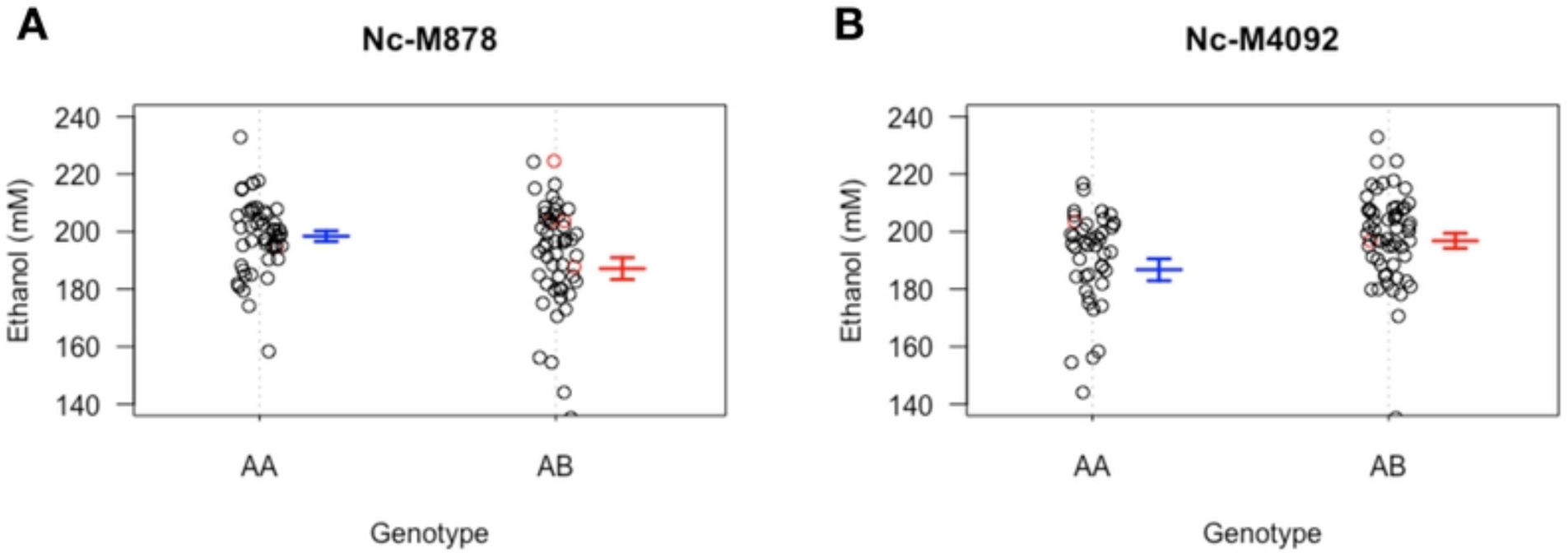
A
Nc-M878
Ethanol (mM)
240
220
200
180
160
140
AA
AB
Genotype
B
Nc-M4092
Ethanol (mM)
240
220
200
180
160
140
AA
AB
Genotype

Figure4

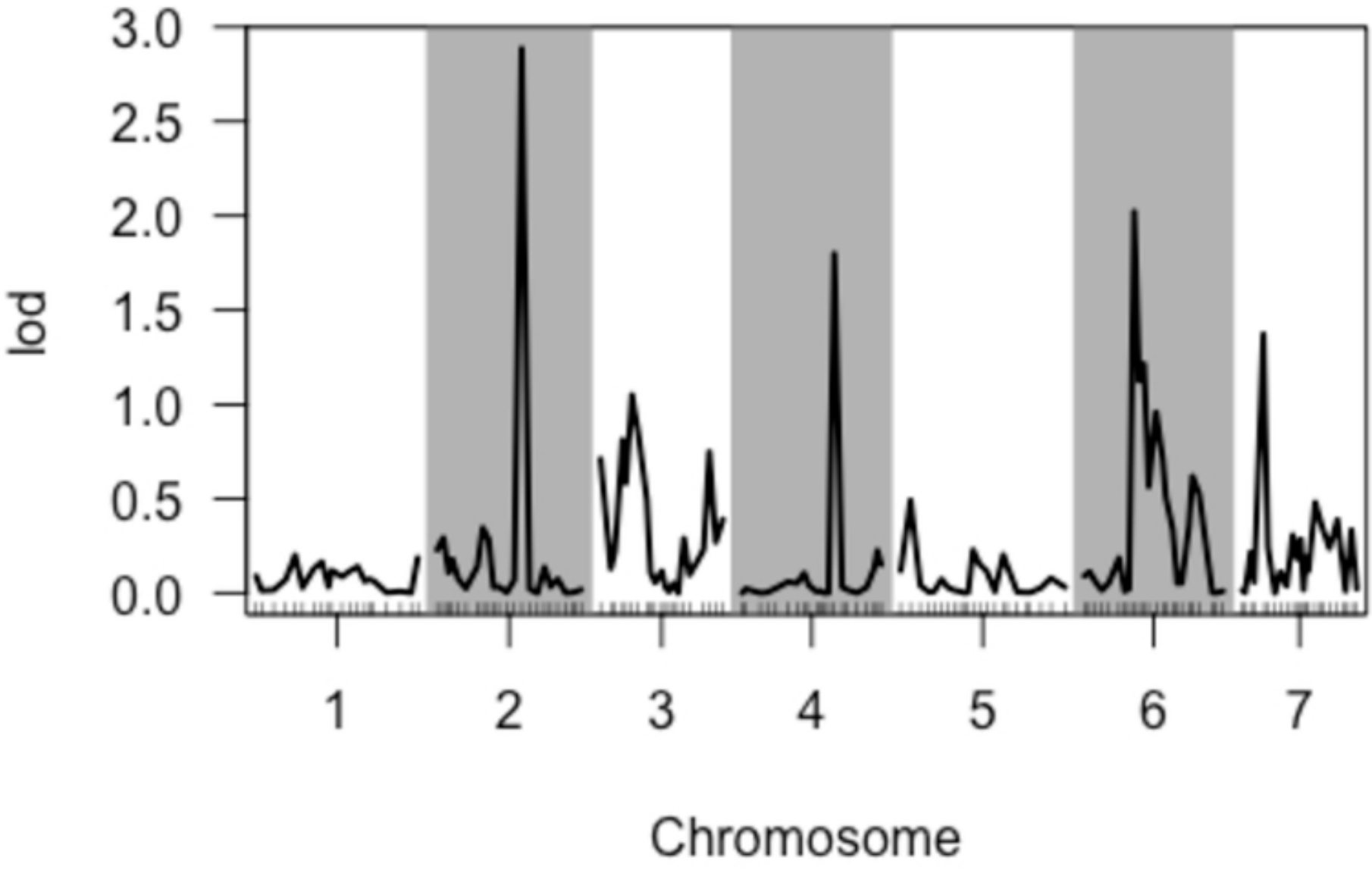

Figure5

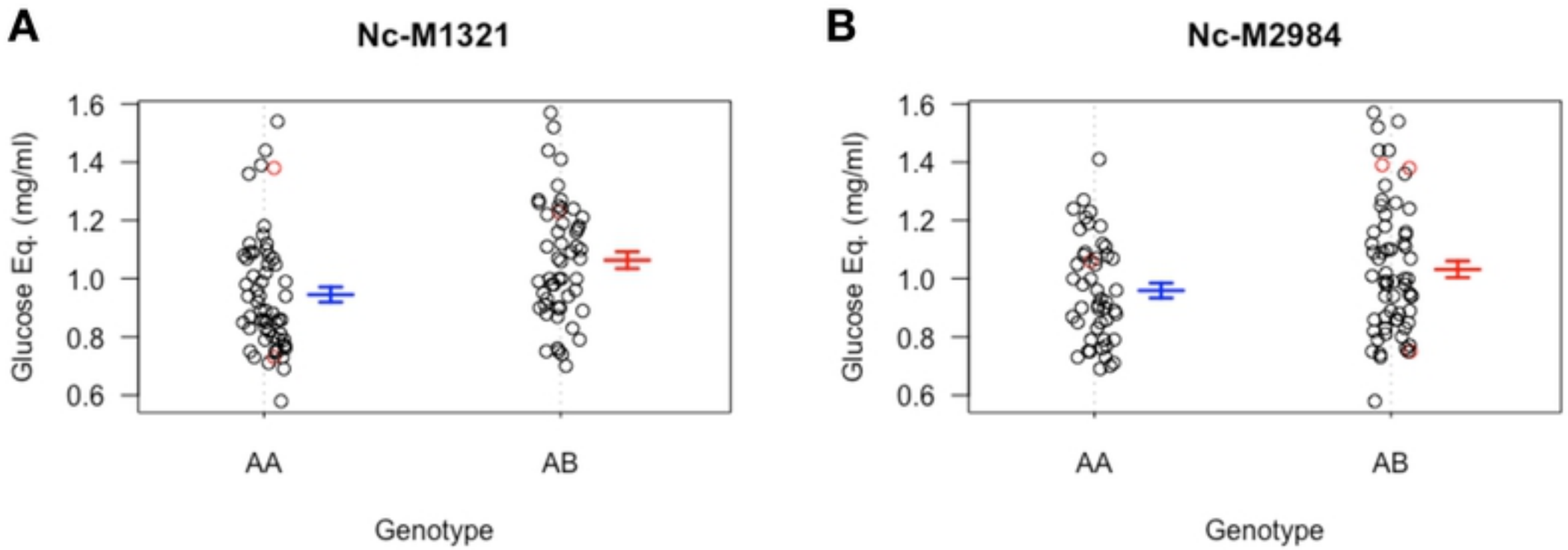
A
Nc-M1321
Glucose Eq. (mg/ml)
1.6
1.4
1.2
1.0
0.8
0.6
AA
AB
Genotype
B
Nc-M2984
Glucose Eq. (mg/ml)
1.6
1.4
1.2
1.0
0.8
0.6
AA
AB
Genotype

Figure6

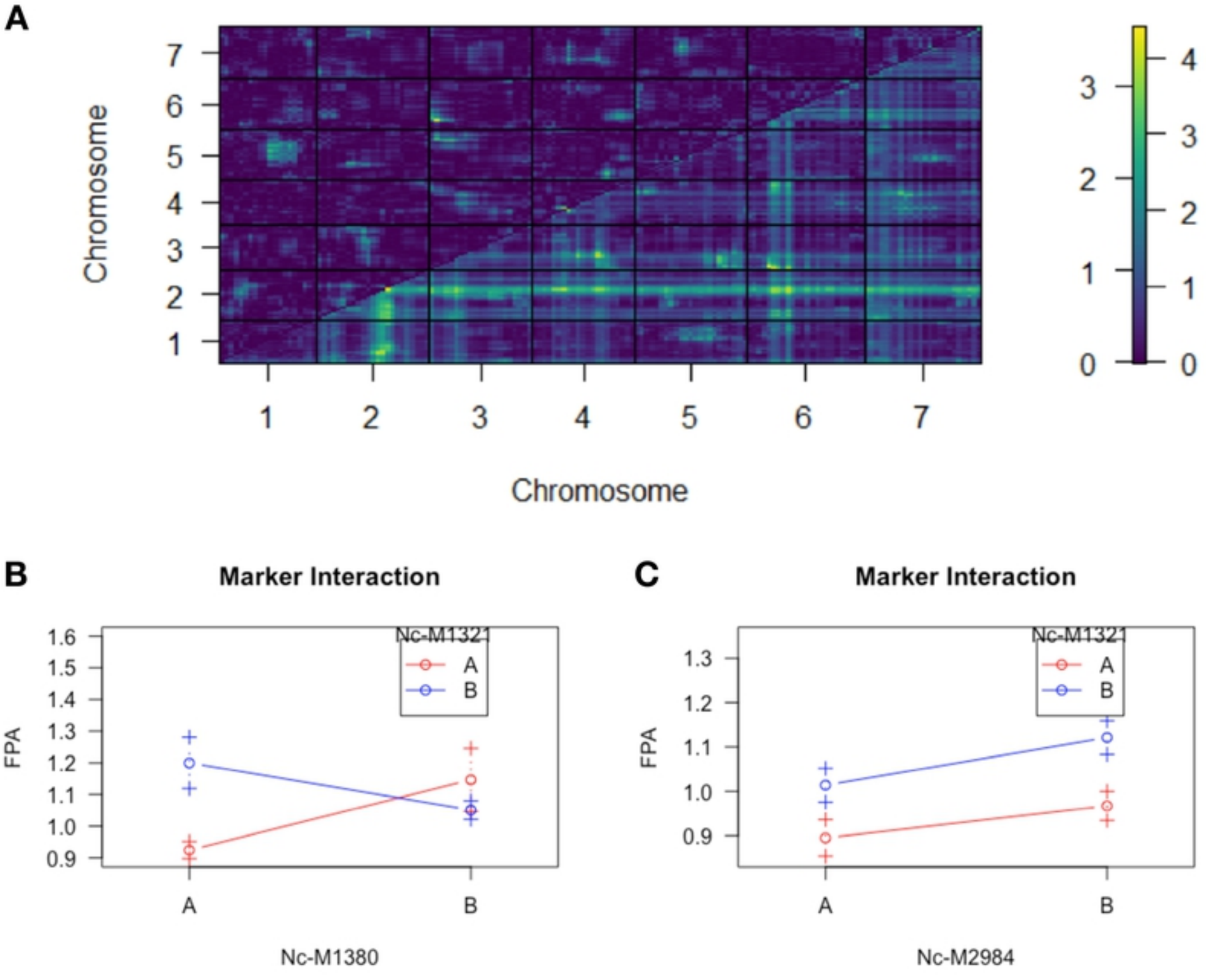
A
Chromosome
7
6
5
4
3
2
1
1
2
3
4
5
6
7
Chromosome
4
3
2
1
0
B
Marker Interaction
Nc-M1321
A
B
FPA
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
A
B
Nc-M1380
C
Marker Interaction
Nc-M1321
A
B
FPA
1.3
1.2
1.1
1.0
0.9
A
B
Nc-M2984

Figure7