# QBio305 - Quantitative Genetics Exercise 6
# Scientific writing

Work by Group 4: Elisaveta Viachaslavovna Lobkis, Florian Tschuschke, Jakob Bohnen, Julia Katharina Ursula Nühlen, and Nicholas Conrad Schmitt

Table of Contents:

# 1. Introduction

## 1.1. The Oregon Wolfe Barley Collection

The Oregon Wolfe Barley (OWB) population is a unique genetic resource that has been extensively utilized in scientific research due to its phenotypic polymorphism and doubled-haploid nature.[1]

The OWB population is characterized by a wide range of phenotypic traits, which have been attributed to its doubled-haploid nature. This means that each cell in the OWB plant has two sets of Chromosomes, which can lead to a high degree of variation in plant traits.
The wide range of phenotypic traits in the OWB population has made it a valuable resource for researchers studying plant genetics and breeding.[2]

The Oregon Wolfe Barleys show several phenotypic traits, such as the qualitative phenotypic trait of two-rowed versus six-rowed barley. This trait refers to the number of grains that develop in each spikelet of the barley head and is a key distinguishing feature of barley types.

Dr. Bob Wolfe - the progenitor of this population - developed the population's parents by systematically crossing recessive alleles into one parent and dominant alleles into the other parent, resulting in dominant and recessive marker stocks.
The overall population comprises 175 barley-doubled haploid offspring resulting from the mating of dominant and recessive marker strains. Notably, 82 were developed using the Hordeum bulbosum (H.b.) method and 93 were developed using another culture (A.C.).[3]

Numerous studies used the OWB population as a subject due to its unique characteristics. For instance, a study by Costa, J., Corey, A., Hayes, P. et al. conducted a molecular mapping of the OWB population. This study provided valuable insights into the genetic structure of the OWB population and helped to clarify its unique genetic characteristics.[4]

The OWB population has been widely used in practical genetic research. For instance, a phenotypically polymorphic barley mapping population was developed using morphological marker stocks as parents. Ninety-four doubled-haploid lines were derived for genetic mapping from an F1 using the Hordeum bulbosum system.

---

[1] Costa, J., Corey, A., Hayes, P., Jobet C., Kleinhofs A., Kopisch-Obusch A., Kramer S., Kudrna D., Li M., Riera-Lizarazu O., Sato K., Szucs P., Toojinda T., Vales M., & Wolfe R. (2001). Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103, 415–424. https://doi.org/10.1007/s001220100622.
[2] Kutlu, I., Çelik, S., Karaduman, Y., & Yorgancılar, Ö. (2023). Phenotypic and genetic diversity of doubled haploid bread wheat population and molecular validation for spike characteristics, end-use quality, and biofortification capacity. PeerJ, 11. https://doi.org/10.7717/peerj.15485.
[3] https://barleyworld.org/owb
[4] Costa, J., Corey, A., Hayes, P., Jobet C., Kleinhofs A., Kopisch-Obusch A., Kramer S., Kudrna D., Li M., Riera-Lizarazu O., Sato K., Szucs P., Toojinda T., Vales M., & Wolfe R. (2001). Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103, 415–424. https://doi.org/10.1007/s001220100622.

This map was constructed using twelve morphological markers, 87 restriction fragment length polymorphism (RFLP), five random amplified polymorphic DNA (RAPD), one sequence-tagged site (STS), one intron fragment length polymorphism (IFLP), 33 simple sequence repeat (SSR), and 586 amplified fragment length polymorphism (AFLP) markers. The genetic map spanned 1,387 cM with an average density of one marker every 1.9 cM.[5]

The OWB population has also been used to study the genetic diversity and population structure of barley. Due to its homozygosity, doubled haploid (DH) wheat like OWB can be used as a variety, a parent in the creation of new hybrids, and a source of genetic mapping studies.[6] This understanding of the genetic diversity of DH wheat lines based on phenotypic traits and DNA polymorphisms is significant in determining strategy in breeding programs.

Such research has showcased the value of the OWB population as a tool for integrating data on quantitative traits with morphological variants. This aids in map-based cloning of genes controlling morphological traits, paving the way for more targeted and effective plant breeding and genetic research.[7]

## 1.2. Linkage maps and mapping qualitative traits

Linkage mapping is a critical tool in plant genetics, providing insights into the distribution of genes in Chromosomes. It builds upon the principle of recombination frequencies between pairs of markers, which are typically genetic markers like Single Nucleotide Polymorphisms (SNPs) or microsatellites.

In the context of qualitative traits, which exist in several discrete states, linkage mapping is particularly useful. For example, in the Oregon Wolfe Barley collection, the nature of the barley as two-rowed or six-rowed is a qualitative trait. Linkage mapping allows us to identify the markers in closest linkage with these traits, information that is pivotal in breeding programs as it enables targeted selection of desired traits.

Linkage maps also play a crucial role in the study of plant genetics, as they can provide valuable insights into the genetic architecture of these traits. This aids in the prediction of phenotypes from genotypes, a crucial step in plant breeding and conservation efforts. While the mapping of qualitative traits involves several steps, the ultimate goal is to infer the location of the trait of interest based on the positions of the associated markers. This understanding is vital for the ongoing study and improvement of the Oregon Wolfe Barley collection, and other plant species.[8]

---

[5] Costa supra note 4

[6] Kutlu, I., Çelik, S., Karaduman, Y., & Yorgancılar, Ö. (2023). Phenotypic and genetic diversity of doubled haploid bread wheat population and molecular validation for spike characteristics, end-use quality, and biofortification capacity. PeerJ, 11. https://doi.org/10.7717/peerj.15485.

[7] Costa, J., Corey, A., Hayes, P., Jobet C., Kleinhofs A., Kopisch-Obusch A., Kramer S., Kudrna D., Li M., Riera-Lizarazu O., Sato K., Szucs P., Toojinda T., Vales M., & Wolfe R. (2001). Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103, 415–424. https://doi.org/10.1007/s001220100622.

[8] Kulwal PL. Trait Mapping Approaches Through Linkage Mapping in Plants. Adv Biochem Eng Biotechnol. 2018;164:53-82. doi: 10.1007/10_2017_49. PMID: 29435601.

## 1.3. Quantitative traits and QTL mapping

Quantitative Trait Loci (QTL) mapping is a statistical approach used to discover the genomic regions associated with variation in quantitative traits. These are traits that are influenced by multiple genes and exhibit a range of phenotypes, such as plant height in the Oregon Wolfe Barley collection.[9]
The process of QTL mapping involves statistically associating the variation in a quantitative trait with the segregation of molecular markers across the genome. This allows to identify QTLs, which are genomic regions that contribute to the genetic variance of the trait.[10]

In the Oregon Wolfe Barley collection context, QTL mapping has been used to study the genetic basis of plant height. Through analysis, several QTLs that contribute to the variation in plant height can be identified, providing valuable insights into the genetic architecture of this trait. The results of QTL mapping can be used in marker-assisted selection, a breeding method that uses molecular markers associated with desirable traits to guide the selection process. This facilitates the development of improved varieties with desired traits.[11]

In summary, QTL mapping provides a valuable tool for understanding the genetic basis of quantitative traits in the Oregon Wolfe Barley collection and can aid in the development of improved barley varieties.

# 2. Aim of the study

This study aims to comprehensively investigate the phenotypic variation within the Oregon Wolfe Barley collection. The analyses encompass essential components, starting with an Analysis of Variance (ANOVA) focused on plant height.

Additionally, the study involves Genetic Linkage Mapping of the qualitative trait two-rowed versus six-rowed barley, identifying the marker in the closest linkage.
Furthermore, a comprehensive Quantitative Trait Loci (QTL) analysis of plant height will be conducted.

# 3. Materials & Methods

## 3.1. Plant Population and Phenotyping

The study focuses on a thorough examination of phenotypic variation within the Oregon Wolfe Barley collection. In order to identify important features, the plant population, which

[9] Doerge, R. (2002). Mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet **3**, 43–52. https://doi.org/10.1038/nrg703.
[10] Doerge, R. (2002). Mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet **3**, 43–52 . https://doi.org/10.1038/nrg703.
[11] Ramalingam, A. P., Mohanavel, W., Kambale, R., Rajagopalan, V. R., Marla, S. R., Prasad, P. V., Muthurajan, R., & Perumal, R. (2023). Pilot-scale genome-wide association mapping in diverse sorghum germplasms identified novel genetic loci linked to major agronomic, root and stomatal traits. Scientific Reports, *13*(1), 1-12. https://doi.org/10.1038/s41598-023-48758-2.

was made up of barley plants that were two months old, underwent thorough phenotyping. The measurements were conducted on five distinct groups, treated as pseudo-replicates for the purpose of this study.

### 3.1.1. Population Size and Traits Scored

The plant population under research consisted of 51 individuals from the Oregon Wolfe Barley collection. This study included measuring several characteristics, such as plant height and spike length, as well as qualitatively examining whether the barley was phenotypically two-rowed or six-rowed.

### 3.1.2. Methods for Phenotyping

Each plant's vertical extension was assessed by measuring its height with a folding rule. Using a 30-cm ruler, the length of the spike was measured. The experiment also involved grading the row type-related qualitative attribute (two- or six-rowed). Careful execution was required during the phenotyping procedure to guarantee precise and repeatable measurements, for that one spike was chosen as a representative for each plant.

## 3.2. Statistics

### 3.2.1. Statistical Methods

To assess the phenotypic variation within the Oregon Wolfe Barley collection, several statistical methods were employed. An Analysis of Variance (ANOVA) with a focus on plant height was used to start the analysis. The next step was doing a genetic linkage mapping analysis to find the markers that were most closely linked to the row type-related qualitative feature.

### 3.2.2. Plant Height ANOVA

An Analysis of Variance (ANOVA)[12] was applied to assess the variation in plant height across the five groups. The pseudo-replicates, treated as true replicates, facilitated the exploration of the impact of different genotypes on plant height. The statistical analysis was conducted using the R programming language and the “aov” function from the “car” package. The analysis further involved the removal of outliers to generate the final ANOVA results.

### 3.2.3. Genetic Linkage Mapping

Genetic linkage mapping was employed to identify markers closely linked to the qualitative trait associated with row type (two-rowed or six-rowed). The R script utilized the “read.cross”, “plotMap”, and “tryallpositions” functions from the genetic mapping package. The analysis provided insights into the genomic regions associated with the observed variations in row type.

[12] Girden, E. R. (1992). *ANOVA: Repeated measures.* Sage Publications, Inc.

### 3.2.4. Quantitative Trait Loci (QTL) Analysis

Quantitative Trait Loci (QTL) analysis[13] was conducted to explore the genetic basis of plant height. The R script included a loop that systematically performed ANOVAs for each marker, extracting p-values and genetic positions. The results were visualized using the “manhattan” function from the “qqman” package, highlighting significant markers. The script for this study used Bonferroni correction to establish significance thresholds.

## 3.3. Software and Data

The statistical analyses were implemented using the R programming language, version 4.3.2 and the software RStudio 2023.09.1 (Build 494). The following R packages were utilized (in brackets we denote the version used):

- ggplot2 (version 3.3.4)
- qqman (version 0.1.9)
- car (version 3.1-2)
- qtl (version 1.66)

The file "dta---ANOVA_QTLanalysis.csv" contains the underlying data, which includes all measurements for all replicates. For the QTL analysis the file “input_map.csv” was used, which contains 2846 markers, as well as its locations on the specific Chromosome.
Finally, we used the file “owb_linakge_mapqualt_phenotypes_WS23.csv” to conduct linkage mapping.

## 3.4. Potential Issues

Despite the careful execution of the analysis, it is imperative to recognise possible limitations in any statistical methodologies. The interpretation of data should take into account issues such as ANOVA’s sensitivity to outliers, the limits of genetic maps for linkage mapping, and the balance between Type I and Type II errors in QTL analysis.
To account for this, we will consult the wider literature for verification of the results of our analysis.

# 4. Results

## 4.1. The ANOVA table

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|---|---|---|---|---|---|
| Genotype | 50 | 59342 | 1186.8 | 2112 | <0.0000000000000002 *** |
| Residuals | 193 | 108 | 0.6 | | |

[13] Miles, C. & Wayne, M. (2008). Quantitative trait locus (QTL) analysis. Nature Education 1(1):208.

Table 1: Results of the ANOVA model

The conducted ANOVA model shows a very significant correlation between the genotype and plant height. With a p-value this low, one could say it is practically impossible that the two are uncorrelated.

## 4.2. Residual Filtering

We then conducted residual filtering. The list of all important residuals can be found on our GitHub repository[14], but these are the ten most correlated residuals:

| marker | chr | pos | pval.allele |
|---|---|---|---|
| M_30248 | 2H | 158.774 | 0.000000000000000000000000000000000000000002814907 |
| M_Zeo1 | 2H | 156.273 | 0.000000000000000000000000000000000000000010691762 |
| M_30106 | 2H | 155.038 | 0.000000000000000000000000000000000000000043332445 |
| M_30396 | 2H | 155.038 | 0.000000000000000000000000000000000000000043332445 |
| M_bPb1505 | 2H | 155.038 | 0.000000000000000000000000000000000000000043332445 |
| M_221983 | 2H | 155.038 | 0.000000000000000000000000000000000000000043332445 |
| M_11227 | 2H | 153.804 | 0.000000000000000000000000000000000000000165678704942 |
| M_FGXOWB00153 | 2H | 151.303 | 0.000000000000000000000000000000000000002804297520971 |
| M_FGXOWB00198 | 2H | 153.804 | 0.000000000000000000000000000000000000003985628496660 |
| M_21181 | 2H | 160.008 | 0.000000000000000000000000000000000000125363466121651 |

Table 2: Ten residual markers with lowest p-value

[14] https://github.com/Nisch206/QBio305-QuantGen

These filtered residuals - we manually remove the outliers mentioned in the lecture - were all located on the second Chromosome.

Below is a Manhattan plot of all the markers - Figure 1 - which shows a significant spike in the area of the second Chromosome.

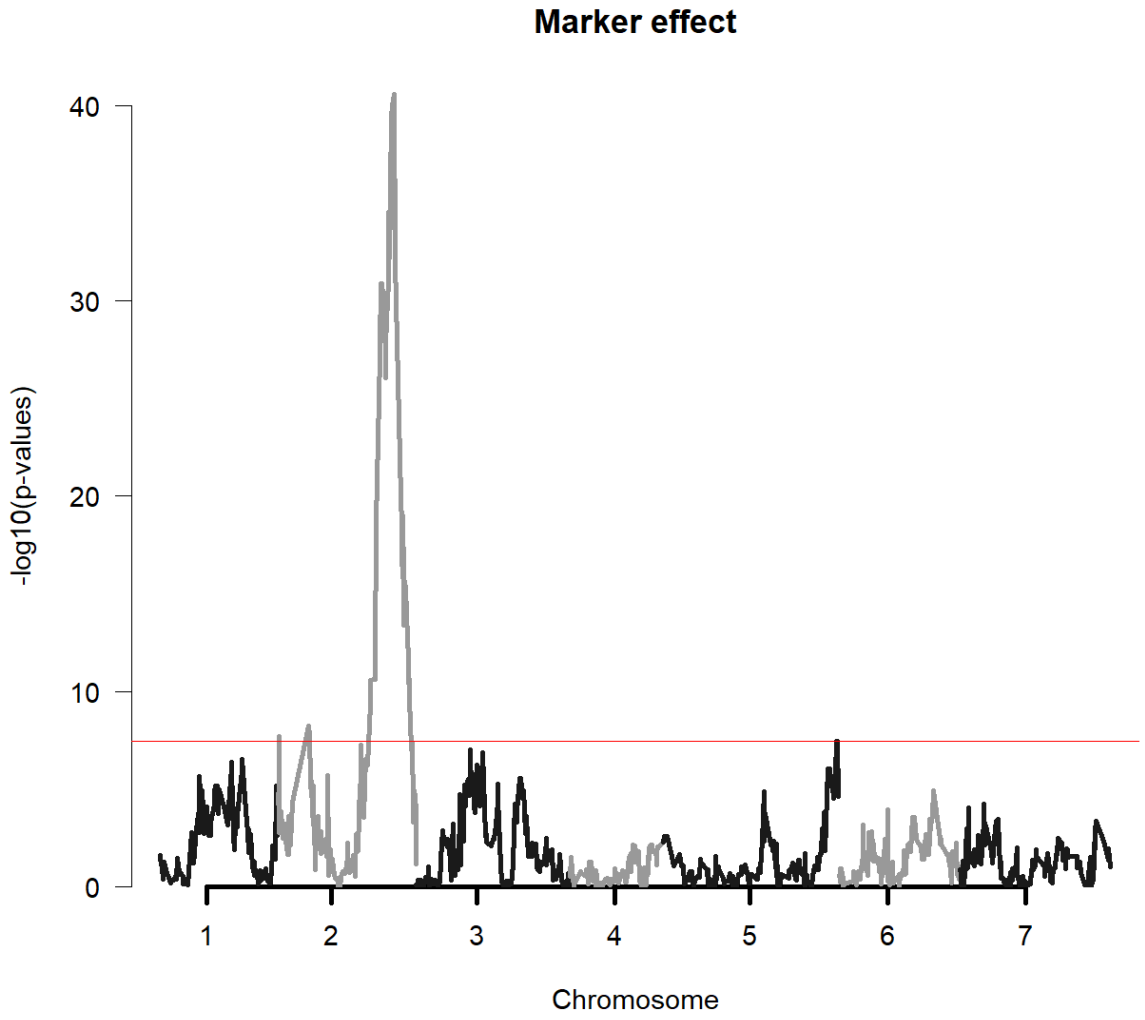


Figure 1: Manhattan Plot of filtered residual markers on provided reference genome

## 4.3. QTL Mapping

Further QTL analysis and linkage analysis produced the following Manhattan plot (Figure 2):

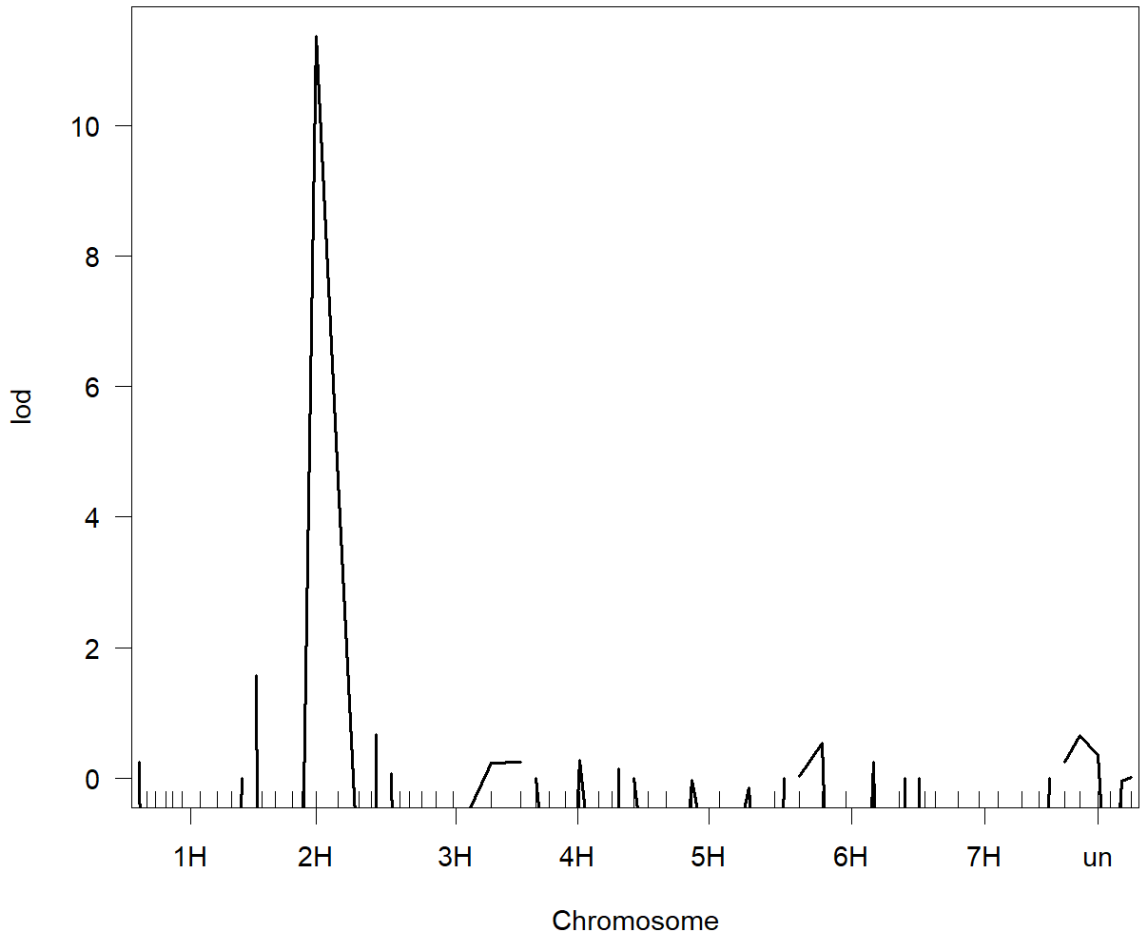


Figure 2: Manhattan Plot of filtered residual markers on provided reference genome after additional analysis

Figure 2 was produced after cross-checking every possible marker location and it can be seen that the most probable location of most quantitative traits is found on the 2H Chromosome, leading to a lod of over 10.
In comparison to the significantly lower lod values of the other Chromosomes, this leads us to believe that 2H is the only relevant Chromosome to study further.

The last plot produced in this study - Figure 3 - is a linkage plot. This plot includes our specific characteristic - 2vs6row - and excludes the characteristics we were not interested in.

Once again, it is observable that our characteristic is located on the second Chromosome.

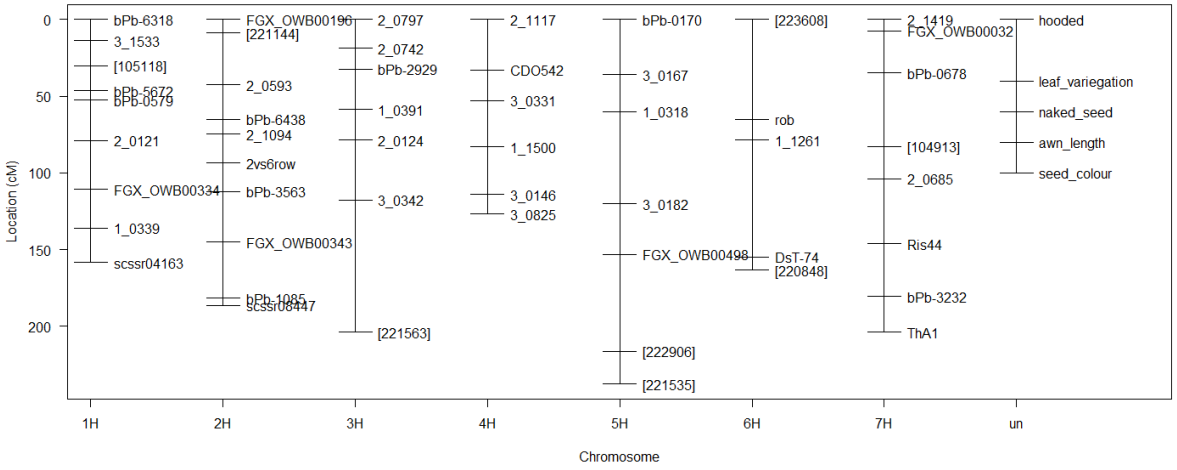


Figure 3: QTL plot of the genome mapping performed by our algorithm

# 5. Discussion

The results displayed in Table 1 concur with our expectations, as it would be unlikely for the plant's height to be independent of its genome.

In Figure 2, we displayed the fact that most of the traits of interest identified were located on the 2H Chromosome. This coincides with other studies which revealed that there are many genes responsible for phenotypic traits on this chromosome[15], including plant height.

In Figure 3, we showed the result of the linkage mapping and that our trait of interest, *2vs6row*, was identified on the 2H Chromosome. This explains the results in Figure 2 and the dominance of the 2H Chromosome as the location of our traits of interest.

Figure 4 presents the map of previously identified genes on the OWB genome in the literature[16]. In both maps, the gene *scssr08447* occurs at the end of the second Chromosome at a distance of around 190 cM[17] to the origin of 2H. This allows us to find out the identity of the gene that is most likely to be responsible for our trait.

We estimated the distance between *scssr08447* and the trait *2vs6row* at around 100cM. By cross-referencing Figure 4, we can identify the gene *VRS1* as being located at the same position on the 2H Chromosome.

---

[15] Chutimanitsakun, Y., Nipper, R.W., Cuesta-Marcos, A. *et al.* (2011). Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics* **12**, 4 . https://doi.org/10.1186/1471-2164-12.

[16] von Korff Schmising, Maria; Chopra-Ufer, Divykriti, Dr.; Hellwig, Timo; Ali, Tahir, Dr. . Genetic linkage mapping of morphological traits in barley. Study course: QBio305: Population and Quantitative Genetics. University of Düsseldorf. Presentation date: 02.11.2023.

[17] Chutimanitsakun, Y., Nipper, R.W., Cuesta-Marcos, A. *et al.* 2011). Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics* **12**, 4. https://doi.org/10.1186/1471-2164-12-4.

This gene was also identified as responsible for this trait in the literature[18], reinforcing our conclusion that this is the gene causing this trait.

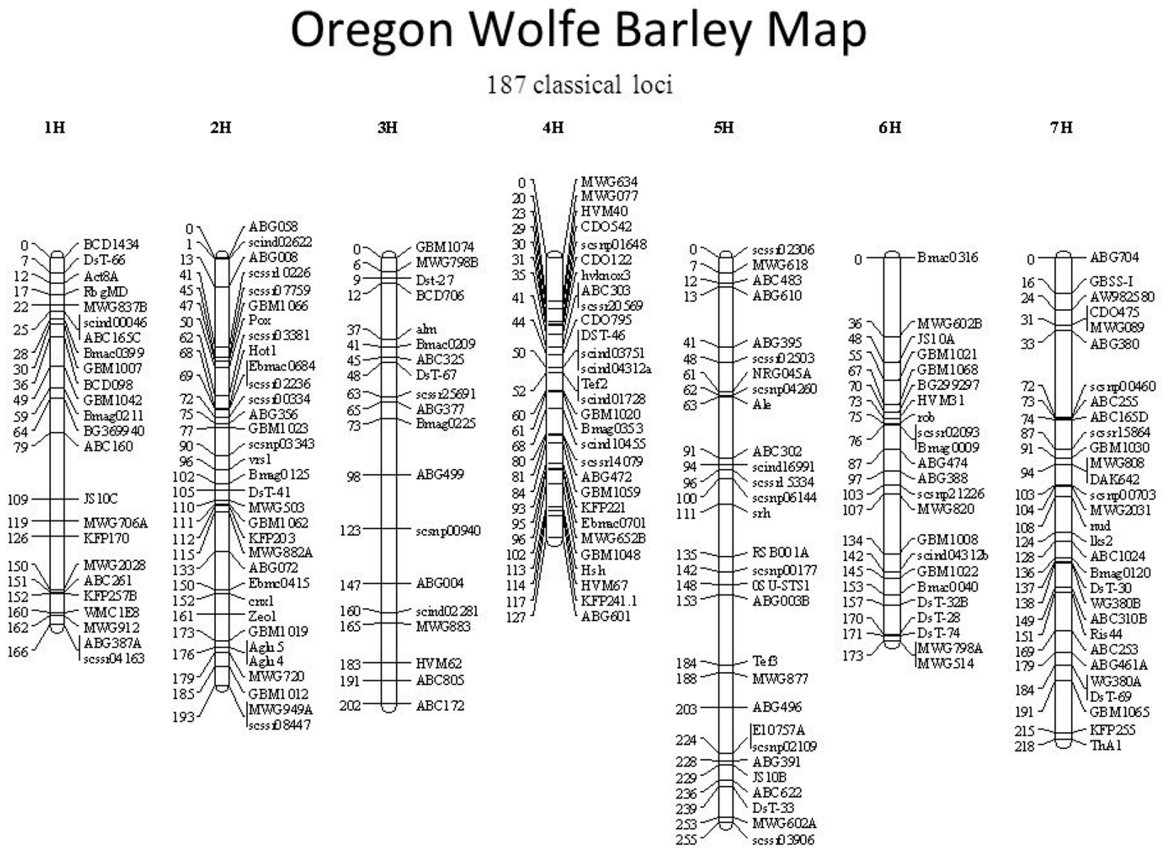


Figure 4: Genomic map of OWB with 187 classical loci on each Chromosome. [19]

This leads us to believe that our analysis was conducted correctly and reached results similar to the literature.

[18] Komatsuda, T., Pourkheiranddish, M., Congfen, H. et. al. 2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences of the United States of America.* 10.1073/pnas.0608580104.

[19] von Korff Schmising, Maria; Chopra-Ufer, Divykriti, Dr.; Hellwig, Timo; Ali, Tahir, Dr. . Genetic linkage mapping of morphological traits in barley. Study course: QBio305: Population and Quantitative Genetics. University of Düsseldorf. Presentation date: 02.11.2023.

# 6. References


1. Costa, J., Corey, A., Hayes, P., Jobet C., Kleinhofs A., Kopisch-Obusch A., Kramer S., Kudrna D., Li M., Riera-Lizarazu O., Sato K., Szucs P., Toojinda T., Vales M., & Wolfe R. (2001). Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103, 415–424. https://doi.org/10.1007/s001220100622.
2. Kutlu, I., Çelik, S., Karaduman, Y., & Yorgancılar, Ö. (2023). Phenotypic and genetic diversity of doubled haploid bread wheat population and molecular validation for spike characteristics, end-use quality, and biofortification capacity. PeerJ, 11. https://doi.org/10.7717/peerj.15485.
3. https://barleyworld.org/owb
4. Kulwal PL. Trait Mapping Approaches Through Linkage Mapping in Plants. Adv Biochem Eng Biotechnol. 2018;164:53-82. doi: 10.1007/10_2017_49. PMID: 29435601.
5. Doerge, R. (2002). Mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet **3**, 43–52. https://doi.org/10.1038/nrg703.
6. Ramalingam, A. P., Mohanavel, W., Kambale, R., Rajagopalan, V. R., Marla, S. R., Prasad, P. V., Muthurajan, R., & Perumal, R. (2023). Pilot-scale genome-wide association mapping in diverse sorghum germplasms identified novel genetic loci linked to major agronomic, root and stomatal traits. Scientific Reports, *13*(1), 1-12. https://doi.org/10.1038/s41598-023-48758-2.
7. Girden, E. R. (1992). *ANOVA: Repeated measures.* Sage Publications, Inc
8. Miles, C. & Wayne, M. (2008). Quantitative trait locus (QTL) analysis. Nature Education 1(1):208.
9. Chutimanitsakun, Y., Nipper, R.W., Cuesta-Marcos, A. *et al.* (2011). Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics* **12**, 4. https://doi.org/10.1186/1471-2164-12.
10. von Korff Schmising, Maria; Chopra-Ufer, Divykriti, Dr.; Hellwig, Timo; Ali, Tahir, Dr. . Genetic linkage mapping of morphological traits in barley. Study course: QBio305: Population and Quantitative Genetics. University of Düsseldorf. Presentation date: 02.11.2023.
11. Komatsuda, T., Pourkheiranddish, M., Congfen, H. et. al. 2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences of the United States of America.* 10.1073/pnas.0608580104.
12. Girden, E. R. (1992). *ANOVA: Repeated measures.* Sage Publications, Inc.
13. Miles, C. & Wayne, M. (2008). Quantitative trait locus (QTL) analysis. Nature Education 1(1):208.
14. https://github.com/Nisch206/QBio305-QuantGen
15. Chutimanitsakun, Y., Nipper, R.W., Cuesta-Marcos, A. *et al.* (2011). Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics* **12**, 4 . https://doi.org/10.1186/1471-2164-12.
16. von Korff Schmising, Maria; Chopra-Ufer, Divykriti, Dr.; Hellwig, Timo; Ali, Tahir, Dr. . Genetic linkage mapping of morphological traits in barley. Study course: QBio305: Population and Quantitative Genetics. University of Düsseldorf. Presentation date: 02.11.2023.
17. Chutimanitsakun, Y., Nipper, R.W., Cuesta-Marcos, A. *et al.* 2011). Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics* **12**, 4. https://doi.org/10.1186/1471-2164-12-4.
18. Komatsuda, T., Pourkheiranddish, M., Congfen, H. et. al. 2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene.

*Proceedings of the National Academy of Sciences of the United States of America.* DOI:10.1073/pnas.0608580104.
19. von Korff Schmising, Maria; Chopra-Ufer, Divykriti, Dr.; Hellwig, Timo; Ali, Tahir, Dr. . Genetic linkage mapping of morphological traits in barley. Study course: QBio305: Population and Quantitative Genetics. University of Düsseldorf. Presentation date: 02.11.2023.

# 7. Supplementary Material

https://github.com/Nisch206/QBio305-QuantGen