Scientific Research Publishing

American Journal of Plant Sciences, 2024, 15, 811-834
https://www.scirp.org/journal/ajps
ISSN Online: 2158-2750
ISSN Print: 2158-2742

Genome-Wide Association Study and Genomic Selection for Plant Growth Habit in Peanuts Using the USDA Public Data

Aurora Manley[1], Madeline Brown[1], Waltram Ravelombola[1,2*], John Cason[2,3], Hanh Pham[4]

[1]Texas A&M AgriLife Research, Vernon, USA
[2]Soil and Crop Sciences, Texas A&M University, College Station, USA
[3]Texas A&M AgriLife Research, Stephenville, USA
[4]Texas A&M AgriLife Research, Lubbock, USA
Email: *waltram.ravelombola@ag.tamu.edu

How to cite this paper: Manley, A., Brown, M., Ravelombola, W., Cason, J. and Pham, H. (2024) Genome-Wide Association Study and Genomic Selection for Plant Growth Habit in Peanuts Using the USDA Public Data. *American Journal of Plant Sciences*, **15**, 811-834.
https://doi.org/10.4236/ajps.2024.159052

Received: April 16, 2024
Accepted: September 23, 2024
Published: September 26, 2024

Abstract

Peanut (*Arachis hypogaea* L.) production is valued at $1.28 billion annually in the USA. Plant growth habit can be used to determine plant population density and cultivation practices a given farmer uses. Erect plants are generally more compact and can be more densely planted unlike plants with more prostrate growth. The objectives of this study were to analyze publicly available datasets to identify single-nucleotide polymorphism (SNP) markers associated with plant growth habit in peanuts and to conduct genomic selection. A genome-wide association study (GWAS) was used to identify SNPs for growth habit type among 775 USDA peanut accessions. A total of 13,306 SNPs were used to conduct GWAS using five statistical models. The models used were single-marker regression, generalized linear model (PCA), generalized linear model (Q), mixed linear model (PCA), and mixed linear model (Q) and a total of 181, 1, 108, 1, and 10 SNPs were found associated with growth habit respectively. Based on this dataset, results showed that genomic selection can achieve up to 61% accuracy, depending on the training population size being used for the prediction. SNP AX-176821681 was found in all models. Gene ontology for this location shows an annotated gene, *Araip.0F3YM,* found 2485 bp upstream of this SNP and encodes for a peptidyl-prolyl cis-trans isomerase. To the best of our knowledge, this is the first report identifying molecular markers linked to plant growth habit type in peanuts. This finding suggests that a molecular marker can be developed to identify specific plant growth habits in peanuts, enabling early generation selection by peanut breeders.

Keywords

SNP, Cultivated Peanut, GS, GWAS, Growth Habit

1. Introduction

Peanuts (*Arachis hypogea* L.) are used for their high oil and protein content and have an annual value of $1.28 billion in the USA [1]. Over 100 countries cultivate peanuts. Peanut consumption provides essential nutritional elements such as folate, copper, potassium, vitamin E, and etc. for human health [2]. Peanut biomass is also a high-quality forage for livestock, providing 700 g of organic matter digestibility and about 140 g of crude protein per kg of dry plant matter [3].

Both peanut morphology and plant growth habit have been used to classify *Arachis hypogaea* L. into two subspecies. The two subspecies are then divided into different botanical varieties [4]. Specifically, plant growth habit is an important trait which affects both agronomic practices and crop yield. Erect plants with small branch angles are more compact; allowing for dense plantings unlike those that are prostrate with big branch angles [5]. There is disagreement regarding if inheritance of the growth habit trait is nuclear or cytoplasmic and if the mechanism controlling branch angle inheritance is polygenic or monogenic [6]-[8]. A chromosomal segment substitution line population was utilized by Fonceka *et al.* (2012) who found that several quantitative trait loci (QTLs) control peanut growth habit [9]. However, Kayam *et al.* (2017) found a major QTL on chromosome B05 for growth habit using bulk segregant analysis with sequencing results [8].

Genome-wide association studies (GWAS) utilize collected phenotype and genotype data from a large sample of unrelated individuals, which was first developed to detect variants among the genetics of human diseases [10]-[12]. Genome-wide single nucleotide polymorphisms (SNPs) identified via array-based genotyping, genotyping-by-sequencing, or resequencing make up the genotype data. GWAS analysis does not require population development and can detect genes with smaller effect sizes, and improve resolution with smaller blocks of linkage disequilibrium (LD) [11] [12].

Statistical methods are used to associate genetic markers with the phenotype being studied. These analysis methods identify SNPs at which variation in genotype is significantly associated with variation in phenotype. Performing an ANOVA on each individual SNP can accomplish this using the hypothesis that there are no differences between the trait mean for any genotype group [13]. Unfortunately, as the number of SNPs used increases, the probability of false positives also increases [13] [14]. Unknown relatedness among individuals is another contributor to false positives. This is because those related individuals form subpopulations within the population. It is difficult to avoid or minimize the unequal relationships within the assembled population for a GWAS study

[15] [16]. If the phenotype is present at a higher frequency in the subpopulation this results in spurious associations with the phenotype. Hence, multiple testing methods are used such as false discovery rate (FDR) and Bonferroni correction [17] [18].

Null markers, unlikely to affect the trait of interest, have been used to estimate population structure effects on test statistics and adjust the final *p* value to reduce false positives [19]. These types of markers were also used to define a set of subpopulations within a dataset; structured association [20]. Once individuals have been grouped in one or more subpopulations, the subpopulation membership is used as a cofactor. The general linear model (GLM) adds the cofactors to correct for population structure [20].

In this study a mixed linear model (MLM) replaced the previously mentioned methods and uses population structure (Q) and kinship (K) to account for relatedness [15]. The kinship matrix uses genotype data from all individuals to estimate the relatedness among them. Using allele frequencies and identity-by-state to estimate identity-by-descent and kinship coefficients is a method [21]-[23]. With the MLM model, false positives are controlled by having a fixed effect of population structure and a random effect of polygenic background which is defined by kinship [15].

Additional models can be used to facilitate identifying the SNPs which are closely associated with the phenotype of interest. Once SNPs have been identified, the location of the gene(s) which control the observed phenotype can be determined. Li *et al.* (2022) used GWAS and bulked segregant analysis to identify loci which control growth-habit related traits among a group of 103 accessions of the U.S. mini-core collection [5]. However, more studies are needed to better understand the genetics of plant growth habit in peanuts. The objective of this study was to conduct a genome-wide association study for plant growth habit in peanuts using the available United State Department of Agriculture public data.

2. Materials and Methods

2.1. Plant Materials and Phenotyping

A total of 775 USDA peanut accessions were phenotyped for growth habit using a binary score 1: spreading, score 2: bunch, and score 3: erect, with the data obtained from the USDA GRIN public data available at https://npgsweb.ars-grin.gov/gringlobal/search.

2.2. Genotyping, Population Structure, Genome-Wide Association Study, and Candidate Gene Search

The Arachis_Axiom2 SNP array was used to genotype the peanut accessions, and this data was made available at https://agdatacommons.nal.usda.gov/ [24]. In the study a total of 13,306 SNPs were used to genotype theaccessions. STRUCTURE 2.3.4 was used to conduct population structure analysis [20]. A total of 10 independent runs, with the Markov Chain Monte Carlo (MCMC) length of burn-in period set to 50,000 and the number of MCMC iterations was 50,000, were conducted to infer population structure (K). The algorithm, developed by Evanno *et al.* (2005), which STRUCTURE Harvester was established from, was used to identify the optimal K value [25] [26]. Each genotype was assigned to a Q-group using a Q-matrix which contained K-vectors corresponding to the optimal K value with a cutoff probability of 0.55. STURCTURE PLOT in STRUCTURE 2.3.4 and the option "Sort by Q" was used to analyze the population structure [20].

Prior to conducting GWAS, SNPs were filtered based on the following criteria: heterozygosity <10%, missing data <10%, and minor allele frequency >5%. After filtering, a total of 13,306 SNPs were used for GWAS analysis in TASSEL 5 [27]. A total of five GWAS models were used for this analysis. The first model used was a single-marker regression model (SMR). The second model was the generalized linear model with principal component (PCA) was added as a covariate (GLM_PCA). The third model used was the generalized linear model with the Q matrix, from the population structure analysis, added as a covariate (GLM_Q). A mixed linear model with Kinship (K) was added to the GLM_PCA model (MLM_PCA+K) as the fourth model. For the last model, a mixed linear model with population stratification being controlled by the Q matrix and Kinship (K) (MLM_Q+K) was used. TASSEL 5's in-built functions were used to estimate both K and PCs. A LOD threshold of > 3.5 was used to identify SNPs significantly associated with plant growth habit phenotype [28]. A candidate gene search was then conducted within a 10-kb region containing a significant SNP and conducted using Peanut Base (https://www.peanutbase.org/taxa/arachis/).

2.3. Genomic-Estimated Breeding Values (GEBVs) and Genomic Selection-Accuracy Assessment

A ridge regression best linear unbiased predictor (rrBLUP) model was then used to compute the genomic-estimated breeding values (GEBVs) [29]. The package "rrBLUP" was used to run the model in R, and the rrBLUP equation was:

$$\gamma = \mathrm{WG}\beta + \mathcal{E}$$

where the y vector phenotype, W the incidence matrix relating the genotype to the phenotype, G the genetic matrix, B the marker effect with $\beta \sim N\left(0, I\sigma z_{\beta}^{2}\right)$, and e the random error. The solution of this equation was:

$$\beta\wedge = (\mathrm{Z}^{\mathrm{T}}\mathrm{Z} + \mathrm{I}\lambda)^{-1}1\mathrm{Z}^{\mathrm{T}}\mathrm{y}$$

where Z = WG. The ridge parameter was defined as $\lambda = \sigma_{e}^{2} / \sigma_{\beta}^{2}$, with σ_{β}^{2} as the marker-effect variance and σ_{e}^{2} as the residual variance.

The effect of the training population size on genomic selection accuracy was evaluated using a 2-fold, 3-fold, 4-fold, 5-fold, and 6-fold cross-validation which corresponded to population sizes of 388, 517, 581, 620, and 646 individuals. A total of 100 replications were used for each cross-validation. Then the accuracy of genomic selection was assessed using Pearson's correlation coefficient between the observed phenotypes in the population and GEBVs [30].

3. Results

3.1. Single Marker Regression (SMR)

Table 1 shows the significant SNPs associated with peanut growth habit using the SMR model. Results from the SMR model indicated that there were 181 SNPs significant to peanut growth habit. These SNPs were located on chromosome A01-A10 and B01-B10, with the majority found on chromosomes B06 and B04. LOD values ranged from 5.43 to 7.65, and the R-square values ranged from 2.74% to 4.46%, indicating that plant growth habit can be controlled by multiple genes with minor effects. The top 10 SNPs with the highest LOD values were AX-147227941 (LOD = 6.9, R^2 = 4.0%), AX-176811670 (LOD = 6.9, R^2 = 4.0%), AX-176822503 (LOD = 6.9, R^2 = 3.6%), AX-176822914 (LOD = 7.0, R^2 = 3.6%), AX-176820260 (LOD = 7.1, R^2 = 4.1%), AX-176820577 (LOD = 7.2, R^2 = 4.2%), AX-176808560 (LOD = 7.2, R^2 = 3.7%), AX-176823020 (LOD = 7.3, R^2 = 4.2%), AX-176821681 (LOD = 7.6, R^2 = 4.0%), and AX-176806956 (LOD = 7.7, R^2 = 4.5%). These SNPs are located on chromosomes A07 (18,840,166 bp), A06 (95,180,807 bp), B06 (29,039,075 bp), B06 (39,793,516 bp), B06 (108,518,747 bp), B06 (129,766,082 bp), B03 (53,474,187 bp), B03 (122,608,001 bp), B06 (122,581,860 bp), and A06 (96,492,176 bp), respectively (**Table 1**).

Table 1. List of SNP markers associated with growth habit in peanuts using different models.

GWAS_models	SNP	Chromosome	Position (bp)	LOD	R^2 (%)
Single Marker Regression	AX-147207638	A01	231,463	5.5	3.3
	AX-176794931	A01	2,479,559	5.6	3.3
	AX-147208618	A01	4,679,934	5.5	3.2
	AX-176811779	A01	96,556,402	5.6	3.3
	AX-147212768	A02	2,937,660	5.9	3.5
	AX-176809518	A02	54,392,732	5.6	3.3
	AX-176814138	A02	69,522,218	5.9	3.4
	AX-176799357	A02	76,275,147	6.7	3.9
	AX-176814184	A02	80,367,821	5.5	3.2
	AX-176802161	A02	88,588,260	5.5	3.2
	AX-176815434	A02	93,648,223	6.2	3.6
	AX-147215334	A03	1,008,880	6.0	3.5
	AX-176795390	A03	32,267,174	5.5	3.2
	AX-147217771	A03	121,816,921	5.8	2.9
	AX-147218175	A03	128,328,312	5.8	3.4
	AX-147218177	A03	128,328,611	5.9	3.5
	AX-147218726	A03	134,513,642	6.0	3.1

Continued

Single Marker Regression	AX-176802330	A04	23,929,426	6.3	3.2
	AX-147219869	A04	33,608,076	6.2	3.6
	AX-176795482	A04	40,982,318	5.5	3.2
	AX-176814816	A04	53,108,353	5.5	3.2
	AX-147220151	A04	76,469,234	5.7	3.4
	AX-147220157	A04	76,790,192	5.8	3.4
	AX-147220160	A04	76,792,848	5.8	3.4
	AX-147220161	A04	76,877,659	5.5	3.2
	AX-147220163	A04	77,270,553	5.6	3.3
	AX-147220164	A04	77,271,279	5.5	3.2
	AX-147220178	A04	78,282,221	5.6	3.3
	AX-147220181	A04	78,282,956	5.5	3.2
	AX-147220186	A04	78,283,551	5.7	3.4
	AX-176792134	A04	78,283,966	5.6	3.3
	AX-176801575	A04	78,395,478	6.4	3.7
	AX-147220191	A04	79,285,408	5.5	3.2
	AX-147220195	A04	79,638,203	6.2	3.6
	AX-147220197	A04	80,024,591	5.8	3.4
	AX-147220198	A04	80,638,497	5.9	3.5
	AX-147220204	A04	81,064,626	6.0	3.5
	AX-176794930	A04	81,067,724	6.3	3.7
	AX-147220210	A04	81,397,986	5.5	3.2
	AX-147220214	A04	82,029,524	6.0	3.5
	AX-147220222	A04	82,272,796	5.9	3.5
	AX-147220225	A04	82,451,216	6.0	3.5
	AX-147220235	A04	83,252,974	6.0	3.5
	AX-176810634	A04	90,906,795	6.2	3.6
	AX-176818527	A04	101,124,999	5.5	3.3
	AX-147248062	A04	104,619,670	5.5	3.2
	AX-176808360	A04	118,900,581	5.6	3.3
	AX-147221136	A04	119,296,024	5.4	3.2
	AX-147223211	A05	91,855,042	5.4	3.2
	AX-176801332	A05	92,527,458	5.5	3.2
	AX-147223291	A05	94,152,944	5.9	3.0
	AX-176799087	A05	101,330,393	5.8	3.4
	AX-147223546	A05	101,331,168	5.8	3.4
	AX-176807288	A06	1,982,008	6.1	3.6

Continued

Single Marker Regression	AX-176807751	A06	2,763,696	6.2	3.6
	AX-176805464	A06	3,491,388	5.7	3.3
	AX-176806115	A06	4,493,236	6.2	3.7
	AX-176819343	A06	6,960,258	5.8	2.9
	AX-176798197	A06	16,318,885	6.1	3.6
	AX-176805389	A06	31,663,747	6.6	3.8
	AX-176811670	A06	95,180,807	6.9	4.0
	AX-176806956	A06	96,492,176	7.7	4.5
	AX-147226321	A06	105,425,114	6.3	3.7
	AX-176803972	A07	4,062,405	5.5	3.2
	AX-147254806	A07	6,489,131	6.1	3.1
	AX-177638761	A07	7,143,266	5.4	3.2
	AX-147227941	A07	18,840,166	6.9	4.0
	AX-147227943	A07	18,840,306	6.0	3.5
	AX-176794097	A07	23,213,115	5.6	2.8
	AX-176792467	A08	1,732,429	6.6	3.9
	AX-177641461	A08	8,340,010	5.5	2.8
	AX-147230402	A08	24,455,849	5.7	3.3
	AX-147230403	A08	24,455,877	5.9	3.5
	AX-176815394	A08	46,505,791	5.8	2.9
	AX-147231998	A08	48,974,813	5.5	2.8
	AX-147233030	A09	19,151,716	6.1	3.5
	AX-147233034	A09	19,152,263	6.4	3.3
	AX-176797333	A09	20,677,308	5.7	3.4
	AX-176795424	A09	87,680,085	5.6	3.3
	AX-147235074	A10	3,387,632	6.2	3.6
	AX-176815545	A10	67,325,207	5.5	2.8
	AX-147264290	A10	84,919,598	5.6	3.3
	AX-176804084	A10	96,704,596	5.9	3.0
	AX-176802196	A10	100,552,153	5.8	3.4
	AX-147236793	A10	103,589,072	5.5	3.2
	AX-147238152	B01	16,824,177	5.5	3.2
	AX-176824168	B01	129,944,418	5.5	3.2
	AX-176796979	B02	61,364,179	5.5	3.2
	AX-176820720	B02	83,222,769	5.6	3.3
	AX-176808560	B03	53,474,187	7.2	3.7
	AX-176823020	B03	122,608,001	7.3	4.2

Continued

Single Marker Regression	AX-176821735	B04	16,211,778	5.6	2.8
	AX-176823480	B04	19,520,056	5.9	3.0
	AX-176822544	B04	27,486,400	6.2	3.6
	AX-176821529	B04	75,330,522	5.8	3.4
	AX-147247704	B04	78,493,772	5.7	3.3
	AX-147247734	B04	82,325,476	5.8	3.4
	AX-147247737	B04	82,843,790	5.5	3.2
	AX-147247739	B04	82,846,208	5.9	3.5
	AX-147247740	B04	82,847,279	6.0	3.5
	AX-147247742	B04	83,847,194	5.5	3.2
	AX-147247744	B04	85,231,477	5.8	3.4
	AX-147247746	B04	85,261,373	6.2	3.6
	AX-147247748	B04	86,383,208	5.8	3.4
	AX-176811428	B04	87,059,163	6.1	3.6
	AX-147247750	B04	87,330,221	5.6	3.3
	AX-147247752	B04	88,398,643	5.7	3.4
	AX-147247757	B04	89,248,043	5.8	3.4
	AX-147247761	B04	89,275,438	5.8	3.4
	AX-176791800	B04	96,900,854	5.6	3.3
	AX-176819114	B04	96,903,250	6.3	3.2
	AX-176820215	B04	100,494,295	6.7	3.9
	AX-176823894	B04	103,717,442	6.3	3.7
	AX-176823955	B04	107,283,159	6.2	3.6
	AX-176809313	B04	113,476,152	6.0	3.5
	AX-147248158	B04	115,634,391	5.9	3.4
	AX-176801426	B04	123,159,927	5.6	3.3
	AX-176807388	B04	124,286,278	5.5	2.7
	AX-176821570	B05	130,262,106	5.6	2.8
	AX-176820650	B06	49,613	5.8	3.0
	AX-176819459	B06	1,080,498	5.7	3.3
	AX-147251757	B06	3,180,468	6.2	3.7
	AX-147251899	B06	4,986,179	5.5	3.2
	AX-176798574	B06	6,107,339	5.5	2.8
	AX-176798149	B06	6,762,762	5.8	2.9
	AX-176822704	B06	9,163,091	6.7	3.9
	AX-176823525	B06	11,994,229	5.5	3.2
	AX-176819407	B06	13,057,095	6.4	3.7

Continued

Single Marker Regression	AX-176817527	B06	16,400,617	5.7	3.4
	AX-176795178	B06	16,470,676	5.8	3.4
	AX-147252588	B06	22,822,411	6.2	3.2
	AX-176808697	B06	23,490,357	5.7	3.3
	AX-176822251	B06	24,151,926	6.0	3.1
	AX-176791527	B06	24,429,699	6.0	3.5
	AX-147252688	B06	27,999,213	6.1	3.6
	AX-176822503	B06	29,039,075	6.9	3.6
	AX-176822914	B06	39,793,516	7.0	3.6
	AX-176819708	B06	45,645,879	5.8	3.4
	AX-176824221	B06	65,129,981	6.7	3.9
	AX-147252963	B06	87,676,700	5.6	3.3
	AX-176823191	B06	102,807,076	5.6	3.3
	AX-176806012	B06	107,609,402	5.5	3.2
	AX-176820260	B06	108,518,747	7.1	4.1
	AX-147253348	B06	118,070,034	6.6	3.8
	AX-147253437	B06	121,537,744	6.7	3.4
	AX-176806377	B06	122,321,048	6.8	3.5
	AX-176821681	B06	122,581,860	7.6	4.0
	AX-176820088	B06	123,275,218	5.9	3.5
	AX-176817763	B06	124,102,491	5.9	3.5
	AX-176823541	B06	124,127,763	6.2	3.2
	AX-176808872	B06	124,127,763	5.9	3.4
	AX-176823574	B06	125,992,228	6.8	4.0
	AX-176822130	B06	127,448,206	6.2	3.6
	AX-147253739	B06	128,287,517	5.6	3.3
	AX-176823068	B06	129,570,003	6.1	3.6
	AX-176820577	B06	129,766,082	7.2	4.2
	AX-147254401	B07	1,449,271	5.7	3.3
	AX-177640154	B07	5,225,213	5.6	3.3
	AX-177640156	B07	6,132,665	6.0	3.5
	AX-177638049	B07	9,189,209	5.5	3.2
	AX-176821319	B07	100,308,848	5.5	2.8
	AX-147256082	B07	105,738,821	5.5	2.8
	AX-177639265	B07	110,494,666	5.8	3.4
	AX-147257104	B08	1,880,289	5.8	3.4
	AX-177644329	B08	2,240,041	5.7	3.4

Continued

Single Marker Regression	AX-177644360	B08	117,912,559	5.9	3.4
	AX-177643206	B09	10,735,371	5.9	3.5
	AX-176823357	B09	115,312,168	5.6	3.3
	AX-177637732	B10	35,189,192	5.6	2.9
	AX-177639197	B10	53,567,565	5.5	2.8
	AX-176823701	B10	53,790,724	5.6	2.8
	AX-176821687	B10	108,714,201	5.8	3.4
	AX-177638968	B10	109,488,824	5.5	3.2
	AX-176821864	B10	114,106,756	5.7	3.3
	AX-177640459	B10	119,790,764	5.9	3.0
	AX-177638504	B10	121,949,032	5.7	3.4
	AX-177638497	B10	127,108,018	5.4	3.2
	AX-177637369	B10	127,616,318	5.7	3.4
	AX-176821433	B10	128,094,475	5.6	2.8
	AX-176822190	B10	131,571,730	5.5	3.2
	AX-147237240	B10	134,914,334	6.5	3.8
Generalized Linear Model (PCA)	AX-176821681	B06	122,581,860	7.1	3.6
Generalized Linear Model (Q)	AX-176814138	A02	69,522,218	5.5	3.2
	AX-176799357	A02	76,275,147	7.0	4.1
	AX-176815434	A02	93,648,223	5.7	3.3
	AX-147215334	A03	1,008,880	6.0	3.5
	AX-147218175	A03	128,328,312	5.7	3.3
	AX-147218177	A03	128,328,611	5.9	3.5
	AX-147218726	A03	134,513,642	5.8	2.9
	AX-176802330	A04	23,929,426	6.2	3.2
	AX-147219869	A04	33,608,076	6.3	3.7
	AX-147220151	A04	76,469,234	5.7	3.3
	AX-147220157	A04	76,790,192	5.9	3.5
	AX-147220160	A04	76,792,848	5.6	3.2
	AX-147220164	A04	77,271,279	5.5	3.2
	AX-147220178	A04	78,282,221	5.5	3.2
	AX-147220181	A04	78,282,956	5.4	3.2
	AX-147220186	A04	78,283,551	5.7	3.3
	AX-176792134	A04	78,283,966	5.6	3.3
	AX-176801575	A04	78,395,478	6.3	3.7
	AX-147220195	A04	79,638,203	6.2	3.6
	AX-147220197	A04	80,024,591	5.8	3.4

Continued

Generalized Linear Model (Q)	AX-147220198	A04	80,638,497	6.0	3.5
	AX-147220204	A04	81,064,626	5.9	3.5
	AX-176794930	A04	81,067,724	6.3	3.7
	AX-147220210	A04	81,397,986	5.4	3.2
	AX-147220214	A04	82,029,524	5.8	3.4
	AX-147220222	A04	82,272,796	5.9	3.4
	AX-147220225	A04	82,451,216	5.8	3.4
	AX-147220235	A04	83,252,974	5.8	3.4
	AX-176810634	A04	90,906,795	6.4	3.7
	AX-176808360	A04	118,900,581	5.6	3.2
	AX-147223291	A05	94,152,944	5.5	2.7
	AX-176799087	A05	101,330,393	5.7	3.4
	AX-147223546	A05	101,331,168	5.8	3.4
	AX-176807288	A06	1,982,008	6.2	3.6
	AX-176807751	A06	2,763,696	6.3	3.7
	AX-176806115	A06	4,493,236	6.5	3.8
	AX-176798197	A06	16,318,885	6.0	3.5
	AX-176805389	A06	31,663,747	6.9	4.0
	AX-176811670	A06	95,180,807	7.0	4.1
	AX-176806956	A06	96,492,176	7.9	4.6
	AX-147226321	A06	105,425,114	6.5	3.8
	AX-147254806	A07	6,489,131	6.0	3.0
	AX-147227941	A07	18,840,166	7.0	4.1
	AX-147227943	A07	18,840,306	5.9	3.4
	AX-176792467	A08	1,732,429	6.9	4.0
	AX-147230402	A08	24,455,849	5.6	3.3
	AX-147230403	A08	24,455,877	5.8	3.4
	AX-147233030	A09	19,151,716	6.0	3.5
	AX-147233034	A09	19,152,263	6.3	3.2
	AX-147235074	A10	3,387,632	6.3	3.6
	AX-176804084	A10	96,704,596	5.5	2.8
	AX-176802196	A10	100,552,153	5.6	3.3
	AX-176808560	B03	53,474,187	7.1	3.6
	AX-176823020	B03	122,608,001	6.8	4.0
	AX-176823480	B04	19,520,056	5.5	2.8
	AX-176822544	B04	27,486,400	6.2	3.6
	AX-176821529	B04	75,330,522	5.6	3.3

Continued

Generalized Linear Model (Q)	AX-147247704	B04	78,493,772	5.5	3.2
	AX-147247734	B04	82,325,476	5.8	3.4
	AX-147247739	B04	82,846,208	5.9	3.4
	AX-147247740	B04	82,847,279	5.9	3.4
	AX-147247744	B04	85,231,477	5.7	3.3
	AX-147247746	B04	85,261,373	6.2	3.6
	AX-147247748	B04	86,383,208	5.6	3.3
	AX-176811428	B04	87,059,163	6.1	3.6
	AX-147247750	B04	87,330,221	5.4	3.2
	AX-147247752	B04	88,398,643	5.7	3.3
	AX-147247757	B04	89,248,043	5.7	3.3
	AX-147247761	B04	89,275,438	5.7	3.3
	AX-176791800	B04	96,900,854	5.5	3.2
	AX-176819114	B04	96,903,250	6.2	3.2
	AX-176820215	B04	100,494,295	6.6	3.8
	AX-176823894	B04	103,717,442	6.4	3.7
	AX-176823955	B04	107,283,159	6.3	3.7
	AX-176809313	B04	113,476,152	5.8	3.4
	AX-147248158	B04	115,634,391	5.5	3.2
	AX-176819459	B06	1,080,498	5.5	3.2
	AX-147251757	B06	3,180,468	6.4	3.7
	AX-176822704	B06	9,163,091	7.1	4.1
	AX-176819407	B06	13,057,095	6.6	3.9
	AX-176795178	B06	16,470,676	5.8	3.4
	AX-147252588	B06	22,822,411	6.0	3.0
	AX-176822251	B06	24,151,926	5.7	2.9
	AX-176791527	B06	24,429,699	5.8	3.4
	AX-147252688	B06	27,999,213	5.9	3.5
	AX-176822503	B06	29,039,075	7.1	3.6
	AX-176822914	B06	39,793,516	7.2	3.7
	AX-176824221	B06	65,129,981	7.0	4.1
	AX-176820260	B06	108,518,747	7.0	4.0
	AX-147253348	B06	118,070,034	6.9	4.0
	AX-147253437	B06	121,537,744	6.3	3.2
	AX-176806377	B06	122,321,048	6.5	3.3
	AX-176821681	B06	122,581,860	7.9	4.1
	AX-176820088	B06	123,275,218	5.6	3.2

Continued

Generalized Linear Model (Q)	AX-176823541	B06	124,127,763	5.8	2.9
	AX-176808872	B06	124,127,763	5.5	3.2
	AX-176823574	B06	125,992,228	7.4	4.3
	AX-176822130	B06	127,448,206	6.4	3.7
	AX-147253739	B06	128,287,517	5.5	3.2
	AX-176823068	B06	129,570,003	6.2	3.6
	AX-176820577	B06	129,766,082	7.4	4.3
	AX-147254401	B07	1,449,271	5.5	3.2
	AX-177640156	B07	6,132,665	5.8	3.4
	AX-177644360	B08	117,912,559	5.5	3.2
	AX-177643206	B09	10,735,371	5.7	3.3
	AX-176823357	B09	115,312,168	5.5	3.2
	AX-177640459	B10	119,790,764	5.5	2.8
	AX-147237240	B10	134,914,334	6.4	3.7
Mixed Linear Model (PCA=K)	AX-176821681	B06	122,581,860	5.6	2.9
Mixed Linear Model (Q + K)	AX-176800551	A01	69,524,629	3.1	1.9
	AX-176807751	A06	2,763,696	3.2	1.9
	AX-176806956	A06	96,492,176	3.4	2.0
	AX-147241123	B02	23,631,160	3.4	2.0
	AX-176808560	B03	53,474,187	3.9	1.9
	AX-147253437	B06	121,537,744	3.5	1.7
	AX-176806377	B06	122,321,048	3.8	1.9
	AX-176821681	B06	122,581,860	4.9	2.5
	AX-176813106	B06	13,430,0181	3.6	2.1
	AX-177642631	B08	39,306,016	3.2	1.9

Figure 1(A) and Figure 1(B) show the Manhattan and QQ plot for the SMR model. A total of 85 significant SNPs were found on the A genome, and 96 significant SNPs were found on the B genome. For the A genome, the chromosomes A01, A02, A03, A04, A05, A06, A07, A08, A09, A10 have 4, 7, 6, 31, 5, 10, 6, 6, 2, and 6 SNPs, respectively (Table 1). For the B genome, the chromosomes B01, B02, B03, B04, B05, B06, B07, B08, B09, and B10 have 2, 2, 2, 27, 1, 37, 7, 3, 2, and 13 SNPs, respectively (Figures 1-5, Table 1).

Figure 1(A) shows clusters of significant SNPs located on the following chromosomes: A04 (76,469,234 bp to 119,296,024 bp), A05 (91,855,042 bp to 101,331,168 bp), A06 (1,982,008 bp to 6,960,258 bp and 9,5180,807 bp to 105,425,114 bp), A08 (1,732,429 bp to 24,455,877 bp), A09 (19,151,716 bp to 20,677,308 bp), B04 (16,211,778 bp to 27,486,400 bp, 75,330,522 bp to 96,903,250 bp, and 100,494,295 bp to 124,286,278 bp), B06 (916,391 bp to 45,645,879 bp and

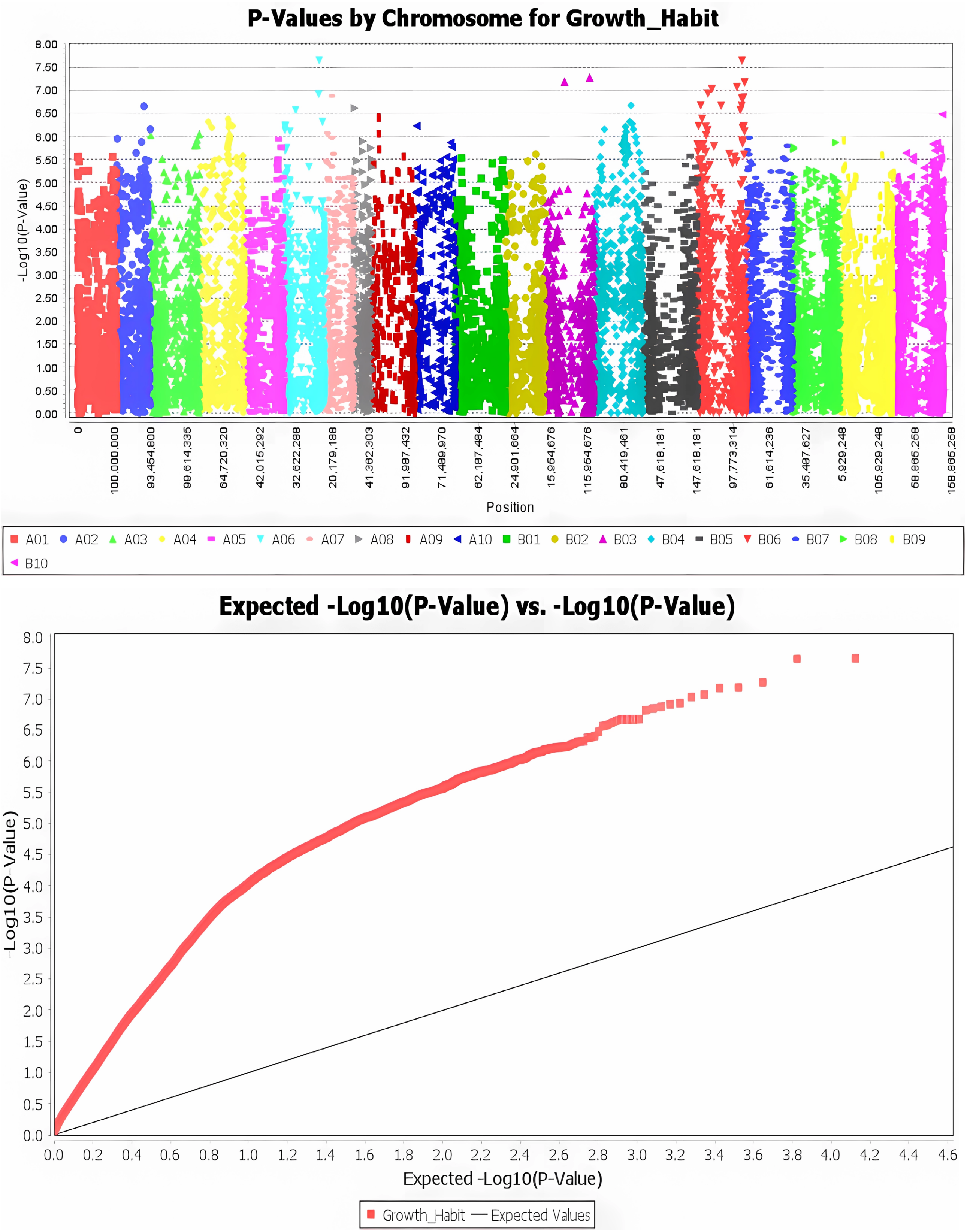

Figure 1. Manhattan and QQ plots using the SMR model.

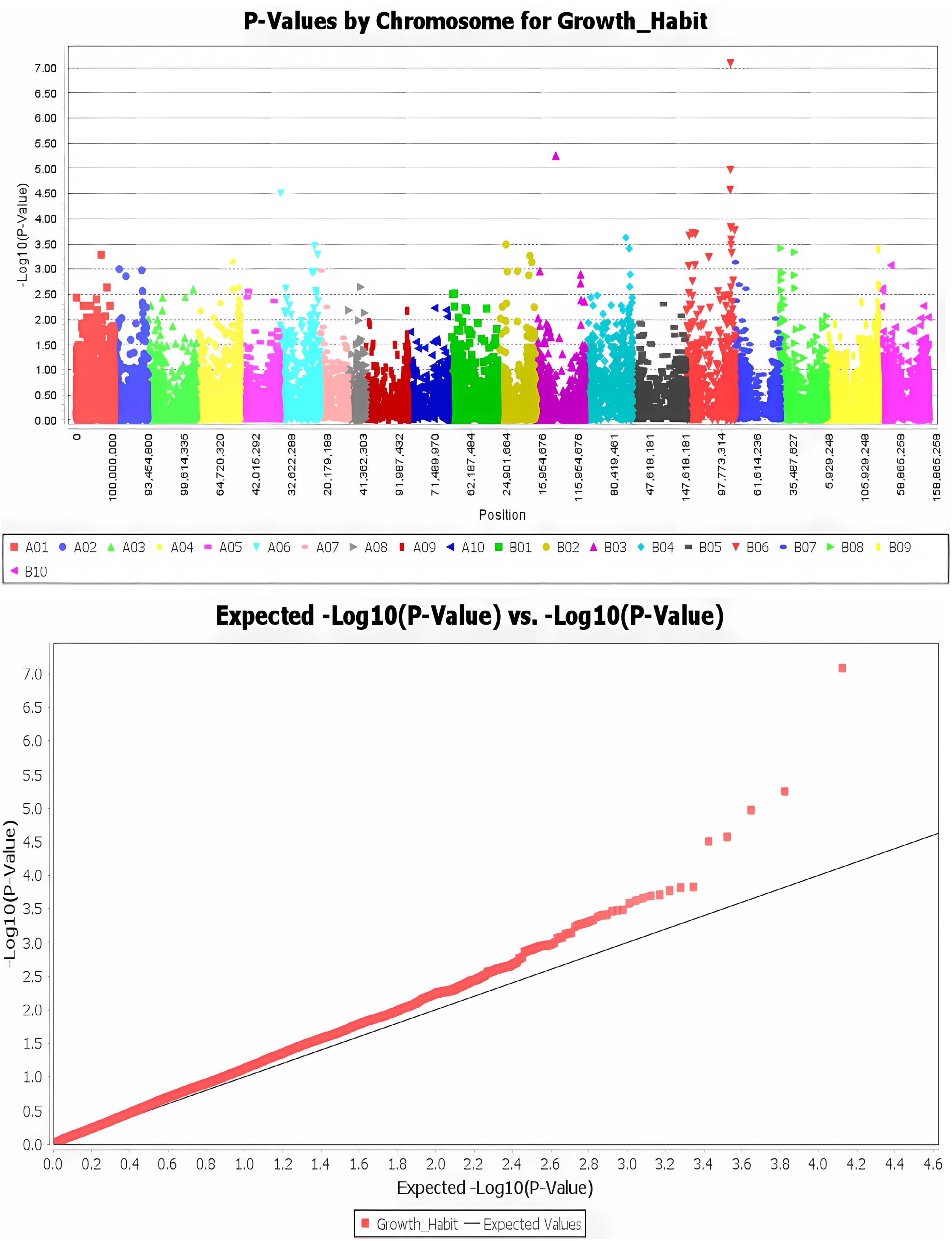

Figure 2. Manhattan and QQ plots using the GLM (PCA) model.

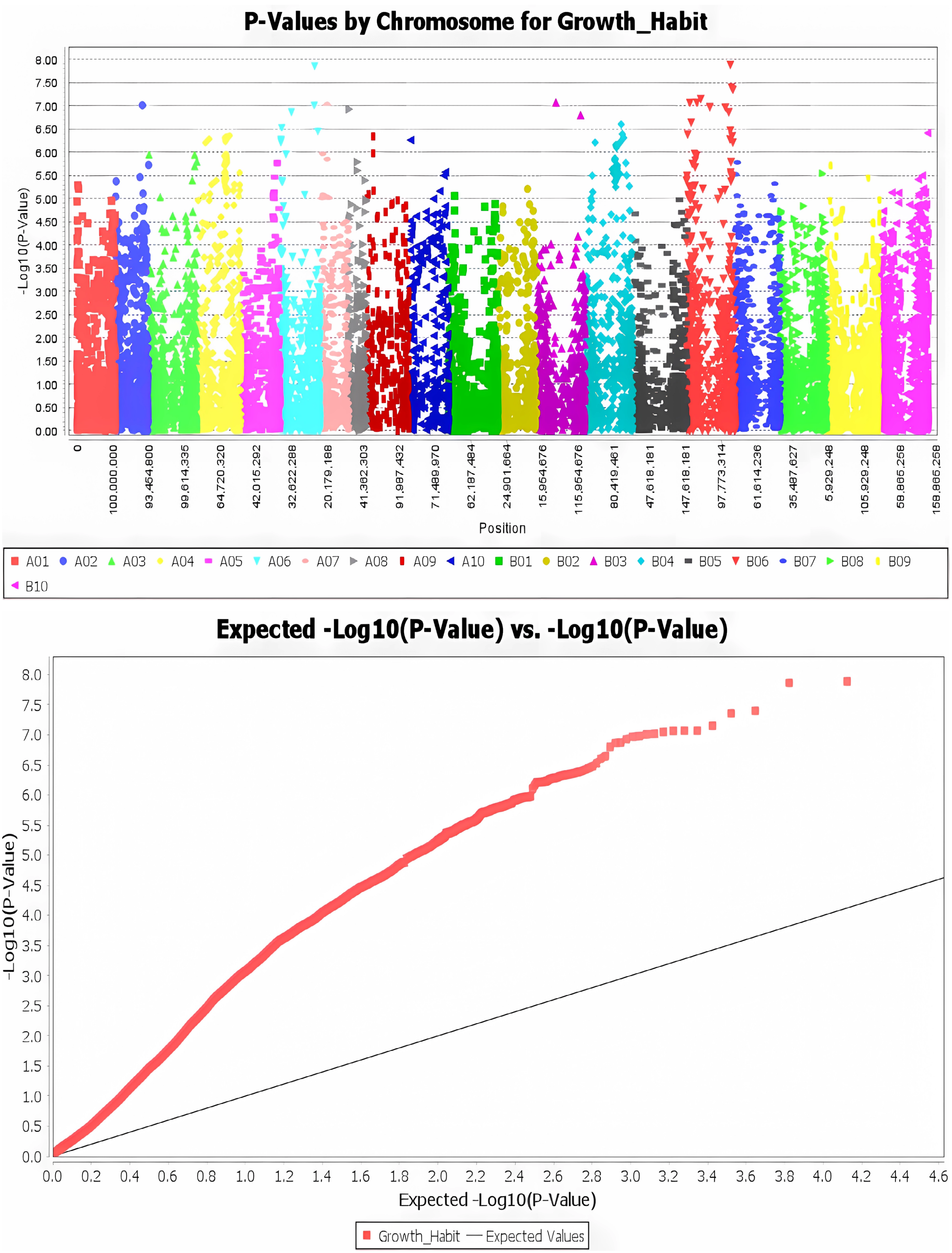

Figure 3. Manhattan and QQ plots using the GLM (Q) model.

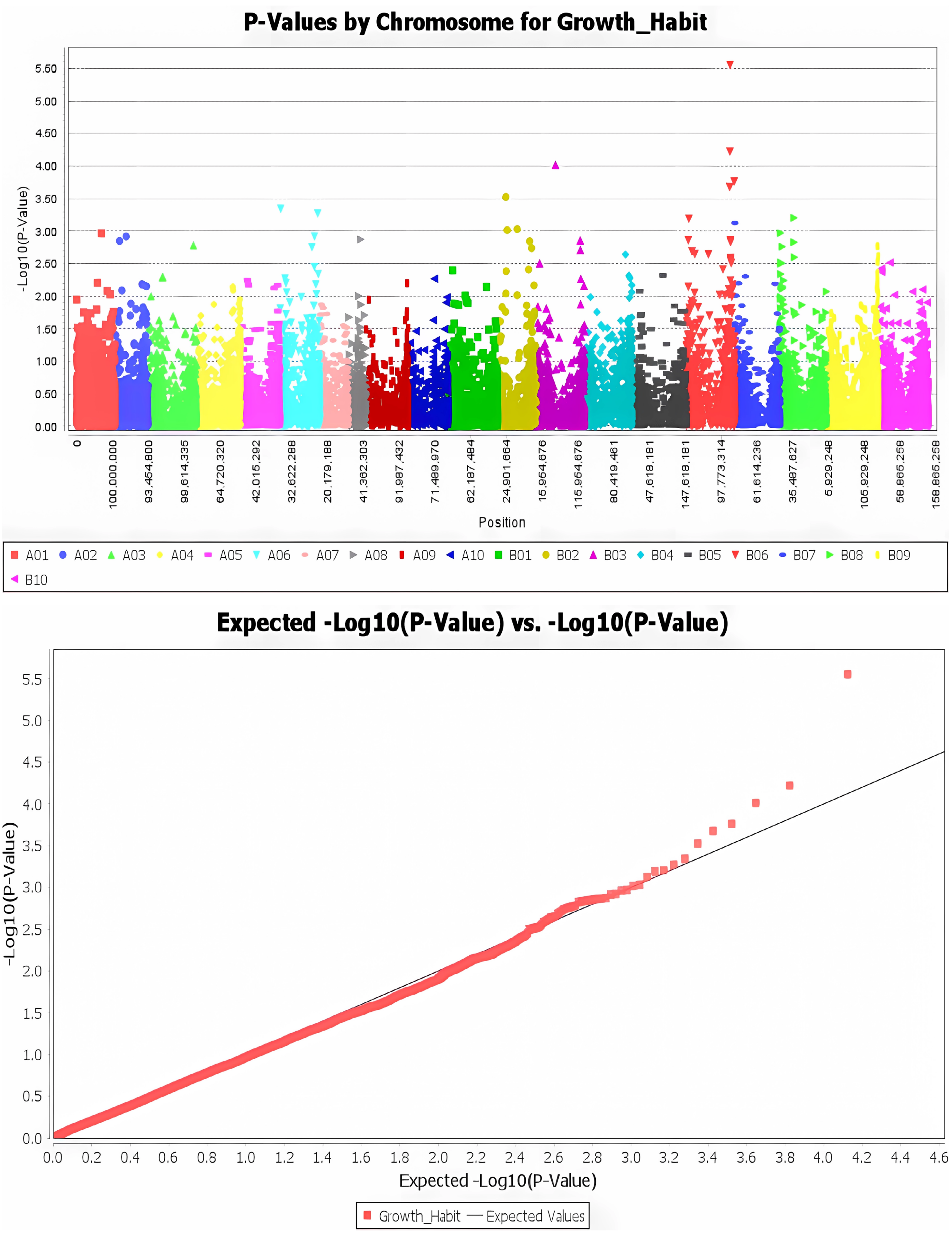

Figure 4. Manhattan and QQ plots using the MLM (PCA + K) model.

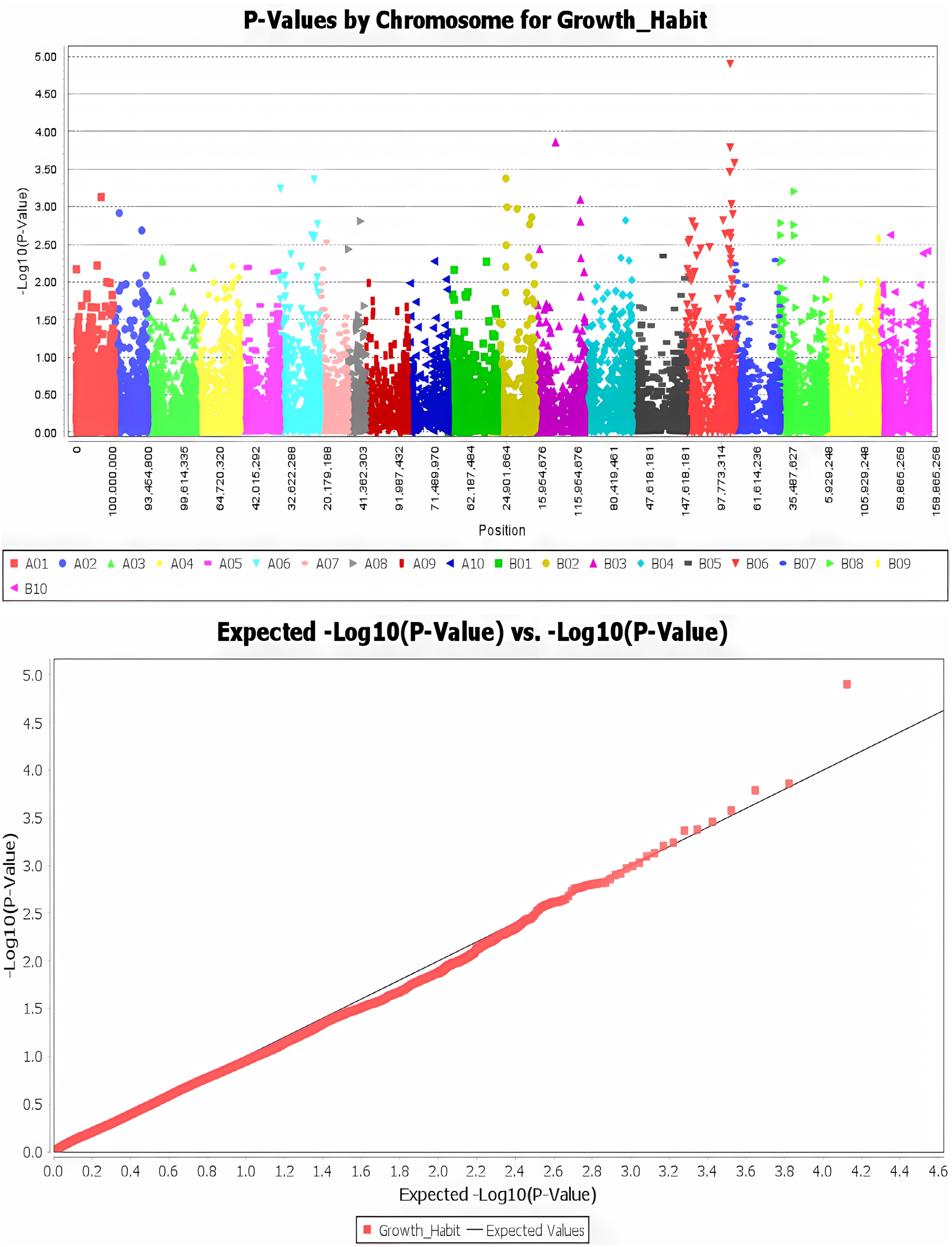

Figure 5. Manhattan and QQ plots using the MLM (Q + K) model.

102,807,076 bp to 129,766,082 bp), and B10 (108,714,204 bp to 134,914,334 bp). The genomic regions where SNP clusters are found suggest that a quantitative trait locus affecting plant growth habit can be found in these areas.

3.2. Generalized Linear Model PCA (GLM_PCA)

Results from the GLM PCA model identified one SNP, AX-176821681, as significantly associated with peanut growth habit. This SNP is located on chromosome B06 (122,581,860 bp), has a LOD score of 7.1, and an R-square value of 3.6%.

3.3. Generalized Linear Model Q (GLM_Q)

The GLM Q model identified 108 SNPs as significantly associated with peanut growth habit. The SNPS were located on chromosomes A02-A10, B03-B04, and B06-B10, most frequently occurring on chromosomes A04, B04, and B06. LOD values ranged from 5.4-7.9, and R-Square values ranged from 2.7% to 4.6%. The ten SNPS with the highest LOD values are AX-176799357 on A02 (76,275,147 bp, LOD = 7.0, and R^2 = 4.1%), AX-14722704 on A07 (18,840,166 bp, LOD = 7.0, and R^2 = 4.1%), AX-176822704 on B06 (9,163,091 bp, LOD = 7.1, and R^2 = 4.1%), AX-176808560 on B03 (53,474,187 bp, LOD = 7.1, and R^2 = 3.6%), AX-17682250 on B06 (29,039,075 bp, LOD = 7.1, and R^2 = 3.6%), AX-176822914 on B06 (39,793,516 bp, LOD = 7.2, and R^2 = 3.7%), AX-176820577 on B06 (129,766,082, LOD = 7.4, and R^2 = 4.3%), AX-176823574 on B06 (125,992,228 bp, LOD = 7.4, and R^2 = 4.3%), AX-176806956 on A06 (96,492,176 bp, LOD = 7.9, and R^2 = 4.6%), and AX-176821681 on B06 (122,581,860 bp, LOD = 7.9, and R^2 = 4.1%). Of the SNPS identified, 52 were located on the A genome and 56 were located on the B genome. For the A genome, the chromosomes A02, A03, A04, A05, A06, A07, A08, A09, A10 had 3, 4, 23, 3, 8, 3, 3, 2, and 3 SNPS respectively. For the B genome, the chromosomes B03, B04, B06, B07, B08, B09, and B10 had 2, 22, 25, 2, 1, 2, and 2 respectively.

Clusters of SNPs significantly associated with plant growth habit were found on the following chromosomes: A03 (128,328,312 bp to 134,513,642 bp), A04 (76,790,192 bp to 118,900,581 bp), A08 (24,455,849 bp to 24,455,877 bp), B04 (75,330,522 bp to 115,634,391 bp), B06 (22,822,411 bp to 29,039,075 bp and 108,518,747 bp to 129,766,082 bp), and B10 (119,790,764 bp to 134,914,334 bp).

3.4. Mixed Linear Model PCA (MLM_PCA + K)

The MLM PCA found one SNP (AX-176821681) to be significantly associated with peanut growth habit on chromosome B06 (122,581,860 bp) with an LOD score of 5.6 and R^2 square of 2.9%.

3.5. Mixed Linear Model Q (MLM_Q + K)

The MLM Q model found 10 SNPs to be significantly associated with peanut growth habit. Located on chromosomes A01 and A06, and B03, B06, and B08, the ten SNPS identified are AX-176800551 on A01 (69,524,629 bp, LOD = 3.1, and R^2 = 1.9%), AX-177642631 B08 (39,306,016 bp, LOD = 3.2, and R^2 = 1.9%), AX-176807751 on A06 (2,763,696 bp, LOD = 3.2, and R^2 = 1.9%), AX-176806956 on A06 (96,492,176 bp, LOD = 3.4, and R^2 = 2.0%), AX-147241123 on B02 (23,631,160 bp, LOD = 3.4, and R^2 = 2.0%), AX-147253437 on B06 (12,153,774 bp) (LOD = 3.5, and R^2 = 1.7%), AX-176813106 B06 (134,300,181 bp) (LOD = 3.6, and R^2 = 2.1%), AX-176806377 B06 (122,321,048 bp, LOD = 3.8, and R^2 = 1.9%), AX-176808560 on B03 (53,474,187 bp, LOD = 3.9, and R^2 = 1.9%), AX-176821681 B06 (122,581,860 bp, LOD = 4.9, and R^2 = 2.5%).

Only a single cluster of SNPs significantly associated with plant growth habit was found on chromosome B06 from 121,537,744 bp to 134,300,181 bp.

3.6. Genomic Selection

Figure 6 shows the genomic selection accuracy of plant growth habit using different training population sizes. The results indicated that a larger training population provided better genomic selection accuracy. The highest accuracy ($r = 0.61$) was obtained for the training population 646, whereas the lowest accuracy was recorded for the training population size 388 ($r = 0.23$). These results demonstrates that genomic prediction can be used as a selection tool for plant growth habits in peanuts.

Figure 6. Accuracy of genomic selection using different training population sizes (388, 517, 581, 620, and 646).

4. Discussion

Plant growth habit in peanuts is used for botanical classification purposes, affects agronomic practices, and overall crop yield. Plants which are erect with small branch angles can be densely planted, unlike those with large branch angles [5]. Researchers disagreed whether inheritance of the growth habit trait is nuclear or cytoplasmic and if branch angle inheritance was under polygenic or monogenic control [6]-[8]. Fonceka *et al.* (2012) used a chromosomal segment substitution line population and found several quantitative trait loci (QTLs) control the growth habit trait in peanuts [9]. However, a bulk segregant analysis with sequencing results revealed a major QTL for the growth habit trait in peanuts on chromosome B05 by Kayam *et al.* [8].

In this study, we used GWAS to identify SNP markers associated with the plant growth habit trait utilizing a publicly available dataset. Significantly associated SNPs were identified in both the A and B sub-genomes. All the SNPs identified had a low R-square value which indicates that plant growth habit is controlled by a small-effect QTL. Previously, GWAS and bulk segregant analysis were used to identify QTL associated with five plant growth habit traits in peanuts. Li *et al.* [5] reported a total of 91 significant SNPs. These SNPs were associated with lateral branch angle (19), main stem height (38), lateral branch height (12), index of plant type (6), and extent radius (16) among the 103 accessions evaluated. These SNPs were distributed among 15 chromosomes, and some were identified for more than one trait. A SNP on chromosome B06 was identified for LBA (lateral branch angle) and ER (extent radius) growth habit traits. These results indicate that chromosome B06 is a good location to identify SNPs related to peanut plant growth habit. Additional research groups have found GWAS to be powerful when seeking which molecular markers are associated with a specific trait of interest [5] [31] [32]. Different GWAS models were tested to identify SNPs which were strongly associated with peanut plant growth habit and could be used to screen future peanut genotypes. A single SNP, AX-176821681, on chromosome B06 was consistent across the five models tested. The single-marker regression and generalized linear models were not as strict as the mixed linear models used. The AX-176821681 SNP had the highest LOD value under each model and overall resulting LOD scores were reduced in the stricter mixed linear models. The closest annotated gene to AX-176821681, using the *Arachis ipaenis* K30076 1.0 data source, was *Araip.0F3YM* (2,485 bp upstream of this SNP) which encodes for a peptidyl-prolyl cis-trans isomerase. Peptidyl-prolyl cis-trans isomerases and foldases catalyze protein isomerization between trans and cis forms of peptide bonds associated with the polypeptide structure by the 180° rotation around the prolyl bond. The isomerase acts as a timer causing protein structure changes to regulate molecular interactions and enzymatic reactions in various pathological and physiological processes [33]. The overexpression of FKBP-like peptidyl-prolyl cis-trans isomerase in *Arabidopsis* could enhance tolerance to drought, ABA, and heat and salt stress [34]. Plants under drought stress have their growth stunted, affecting the plant growth habit of a peanut genotype. Thus, *Araip.0F3YM* could be a candidate gene for plant growth habit in peanuts. Results also suggest that genomic selection can achieve an accuracy of 61% depending on the training population size being used for the prediction. This indicates that genomic selection can be implemented in a peanut breeding program to predict and select plant growth habits. A similar accuracy was found for genomic selection for sting nematode resistance in peanuts

[31] and soybean cyst nematode in soybean [34]. However, the data should be optimized by exploring additional genomic selection models.

5. Conclusion

To the best of our knowledge, this is the first report on identifying molecular markers associated with growth habit in peanut genotypes from the USDA germplasm collection. A total of 181, 1, and 1 SNPs were found associated with growth habit in peanuts using the singe-marker regression, mixed linear model, and generalized linear model, respectively. One SNP was consistently found in all three models, resulting in a molecular marker that can be used to screen for plant growth habit.

Acknowledgements

We are grateful to the USDA for conducting the phenotypic evaluation and to the National Ag Library and Peanut Base for the availability of the genotypic data.

Funding

This work was supported in part by the National and Texas Peanut Producers Board, the USDA National Institute of Food and Agriculture Hatch Project Accession Number 1025956, the USDA National Institute of Food and Agriculture Hatch Project Accession Number 7003209, the Texas A&M Institute for Health and Agriculture, the Texas A&M Advancing Market to Discovery, and the Texas A&M AgriLife Research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] USDA (2020) USDA—National Agricultural Statistics Service. http://www.nass.usda.gov

[2] Kris-Etherton, Zhao, G.X., *et al.* (2001) The Effects of Nuts on Coronary Heart Disease Risk. *Nutritional Reviews*, **59**, 103-111. https://doi.org/10.1111/j.1753-4887.2001.tb06996.x

[3] Blummel, M., Devulapalli, R., Khan, A., Nigam, N., Upadhyaya, H.D. and Vellaikumar, S. (2005) Preliminary Observations on Livestock Productivity in Sheep Fed Exclusively on Haulms from Eleven Cultivars of Groundnut. *International Arachis Newsletter*, **25**, 54-57.

[4] Hayes, T.R. (1933) The Classification of Groundnut Varieties with a Preliminary Note on the Inheritance of Some Characters. *Tropical Agriculture*, **10**, 318-327.

[5] Li, L., Chen, C.Y., Chen, K., Cui, S., Dang, P., Liu, L., Wei, X. and Yang, X. (2022) GWAS and Bulked Segregant Analysis Reveal the Loci Controlling Growth Habit-Related Traits in Cultivated Peanut (*Arachis hypogaea* L.). *BMC Genomics*, **23**, 1-13. https://doi.org/10.1186/s12864-022-08640-3

[6] Patel, J.S., John, C.M. and Seshadri, C.R. (1936) The Inheritance of Characters in the

Groundnut *Arachis hypogaea*. *Indian Academy of Sciences*, **3**, 214-233. https://doi.org/10.1007/BF03047073

[7] Gan, X.M., Cao, Y.L. and Gu, S.Y. (1984) Genetic Variation of Several Quality Traits in Peanut. *Peanut Science and Technology*, **2**, 2.

[8] Kayam, G., Brand, Y., Hedvat, I., *et al.* (2017) Fine Mapping the Branching Habit Trait in Cultivated Peanut by Combining Bulked Segregant Analysis and High-Throughput Sequencing. *Frontiers in Plant Science*, **8**, Article 467.

[9] Fonceka, D., Tossim, H., Rivallan, R., Vignes, H., Lacut, E., de Bellis, F., *et al.* (2012) Construction of Chromosome Segment Substitution Lines in Peanut (*Arachis hypogaea* L.) Using a Wild Synthetic and QTL Mapping for Plant Morphology. *PLOS ONE*, **7**, e48642. https://doi.org/10.1371/journal.pone.0048642

[10] Lander, E.S. (1996) The New Genomics: Global Views of Biology. *Science*, **274**, 536-539. https://doi.org/10.1126/science.274.5287.536

[11] Risch, N. and Merikangas, K. (1996) The Future of Genetic Studies of Complex Human Diseases. *Science*, **273**, 1516-1517. https://doi.org/10.1126/science.273.5281.1516

[12] Lander, E.S. and Schork, N.J. (1994) Genetic Dissection of Complex Traits. *Science*, **265**, 2037-2048. https://doi.org/10.1126/science.8091226

[13] Bush, W.S. and Moore, J.H. (2012) Genome-Wide Association Studies. *PLOS Computational Biology*, **8**, e1002822. https://doi.org/10.1371/journal.pcbi.1002822

[14] Brzyski, D., Peterson, C.B., Sobczyk, P., Candès, E.J., Bogdan, M. and Sabatti, C. (2017) Controlling the Rate of GWAS False Discoveries. *Genetics*, **205**, 61-75. https://doi.org/10.1534/genetics.116.193987

[15] Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., *et al.* (2005) A Unified Mixed-Model Method for Association Mapping That Accounts for Multiple Levels of Relatedness. *Nature Genetics*, **38**, 203-208. https://doi.org/10.1038/ng1702

[16] Zhu, C. and Yu, J. (2009) Nonmetric Multidimensional Scaling Corrects for Population Structure in Association Mapping with Different Sample Types. *Genetics*, **182**, 875-888. https://doi.org/10.1534/genetics.108.098863

[17] Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B*: *Statistical Methodology*, **57**, 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

[18] Storey, J.D. and Tibshirani, R. (2003) Statistical Significance for Genomewide Studies. *Proceedings of the National Academy of Sciences*, **100**, 9440-9445. https://doi.org/10.1073/pnas.1530509100

[19] Devlin, B. and Roeder, K. (1999) Genomic Control for Association Studies. *Biometrics*, **55**, 997-1004. https://doi.org/10.1111/j.0006-341x.1999.00997.x

[20] Pritchard, J.K., Stephens, M. and Donnelly, P. (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945-959. https://doi.org/10.1093/genetics/155.2.945

[21] Loiselle, B.A., Sork, V.L., Nason, J. and Graham, C. (1995) Spatial Genetic Structure of a Tropical Understory Shrub, Psychotria Officinalis (*Rubiaceae*). *American Journal of Botany*, **82**, 1420-1425. https://doi.org/10.1002/j.1537-2197.1995.tb12679.x

[22] VanRaden, P.M. (2008) Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, **91**, 4414-4423. https://doi.org/10.3168/jds.2007-0980

[23] Speed, D. and Balding, D.J. (2014) Relatedness in the Post-Genomic Era: Is It Still Useful? *Nature Reviews Genetics*, **16**, 33-44. https://doi.org/10.1038/nrg3821

[24] Otyama, P.I., Kulkarni, R., Chamberlin, K., Ozias-Akins, P., Chu, Y., Lincoln, L.M., *et al.* (2020) Genotypic Characterization of the U.S. Peanut Core Collection. *G3 Genes|Genomes|Genetics*, **10**, 4013-4026. https://doi.org/10.1534/g3.120.401306

[25] Evanno, G., Regnaut, S. and Goudet, J. (2005) Detecting the Number of Clusters of Individuals Using the Software structure: A Simulation Study. *Molecular Ecology*, **14**, 2611-2620. https://doi.org/10.1111/j.1365-294x.2005.02553.x

[26] Earl, D.A. and vonHoldt, B.M. (2011) Structure Harvester: A Website and Program for Visualizing Structure Output and Implementing the Evanno Method. *Conservation Genetics Resources*, **4**, 359-361. https://doi.org/10.1007/s12686-011-9548-7

[27] Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007) TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics*, **23**, 2633-2635. https://doi.org/10.1093/bioinformatics/btm308

[28] Kaler, A.S., Dhanapal, A.P., Ray, J.D., King, C.A., Fritschi, F.B. and Purcell, L.C. (2017) Genome-Wide Association Mapping of Carbon Isotope and Oxygen Isotope Ratios in Diverse Soybean Genotypes. *Crop Science*, **57**, 3085-3100. https://doi.org/10.2135/cropsci2017.03.0160

[29] Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. (2001) Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, **157**, 1819-1829. https://doi.org/10.1093/genetics/157.4.1819

[30] Qin, J., Shi, A., Song, Q., Li, S., Wang, F., Cao, Y., *et al.* (2019) Genome Wide Association Study and Genomic Selection of Amino Acid Concentrations in Soybean Seeds. *Frontiers in Plant Science*, **10**, Article 1445. https://doi.org/10.3389/fpls.2019.01445

[31] Ravelombola, W., Cason, J., Tallury, S., Manley, A. and Pham, H. (2022) Genome-wide Association Study and Genomic Selection for Sting Nematode Resistance in Peanut Using the USDA Public Data. *Journal of Crop Improvement*, **37**, 273-290. https://doi.org/10.1080/15427528.2022.2087127

[32] Zhou, X., Guo, J., Pandey, M.K., Varshney, R.K., Huang, L., Luo, H., *et al.* (2021) Dissection of the Genetic Basis of Yield-Related Traits in the Chinese Peanut Mini-Core Collection through Genome-Wide Association Studies. *Frontiers in Plant Science*, **12**, Article 637284. https://doi.org/10.3389/fpls.2021.637284

[33] Lu, K.P., Finn, G., Lee, T.H. and Nicholson, L.K. (2007) Prolyl Cis-Trans Isomerization as a Molecular Timer. *Nature Chemical Biology*, **3**, 619-629. https://doi.org/10.1038/nchembio.2007.35

[34] Alavilli, H., Lee, H., Park, M., Yun, D. and Lee, B. (2017) Enhanced Multiple Stress Tolerance in Arabidopsis by Overexpression of the Polar Moss Peptidyl Prolyl Isomerase FKBP12 Gene. *Plant Cell Reports*, **37**, 453-465. https://doi.org/10.1007/s00299-017-2242-9