

Pathway-based analysis enables deeper mining of GWAS data on H/L in chickens

Jie Wang

institute of animal science

Bo Zhu

institute of animal science

Jie Wen

institute of animal science

Qinghe Li

institute of animal science

Guiping Zhao (✉ zhaogpcaas@caas.cn)

Research

Keywords: H/L, GWAS, Pathway

Posted Date: June 4th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-31730/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Heterophils are refers to white blood cells with phagocytosis and killing functions in the avian immune system, These cells identify and kill pathogenic microorganisms through precise regulatory mechanisms. As a simple index, the heterophil/lymphocyte ratio (H/L) in the blood reflects the immune system status of chickens. H/L is a complex trait affected by multiple genetic loci, but single nucleotide polymorphisms (SNPs) significantly associated with traits can usually explain only part of the phenotypic variation. Combining a genome-wide association study (GWAS) with pathway analysis can improve understanding of the biological pathways affecting traits. Our objective was to conduct a SNP- and pathway-based analysis to identify possible biological mechanisms involved in H/L traits.

Methods GWAS for H/L was performed in 1,317 Cobb broilers to identify significant single nucleotide polymorphisms (SNPs) associated with H/L. Eight SNPs ($P < 1/8,068$) reached a significant level of association. The following results were obtained through a series of analyses, including pathway-based association analysis and analysis of the proportion of phenotypic variance explained by SNPs.

Results On the basis of GWAS analysis results, one associated SNP (5% genome-wide significance ($6.80E-6, 0.05/8,068$)) on GGA 1(chicken chromosome 1) and seven suggestively associated SNPs with a trend toward significance($1.24E-4, 1/8,068$) on GGAs 1, 7, 13 and 19 were detected. The significant SNP on GGA 19 was in Complement C1q Binding Protein (C1QBP). The wild-type and mutant individuals showed significant differences in H/L at five loci ($P < 0.05$). According to the results of pathway-based analysis, nine associated pathways ($P < 0.05$) were identified, including small cell lung cancer, proteoglycans in cancer, sulfur relay system, pathways in cancer, gastric acid secretion and purine metabolism. By combining GWAS with pathway analysis, we found that all SNPs after QC explained 12.4% of the phenotypic variation in H/L and 53 SNPs associated with H/L explained as much as 9.7% of the phenotypic variation in H/L.

Conclusions Our findings contribute to understanding of the genetic regulation of H/L and provide theoretical support.

Background

With the continuous increase in intensive agriculture, the poultry industry has developed rapidly. Disease has become a major issue in the poultry industry in China. Vaccination is a common method of disease prevention and treatment, but it has raised concerns regarding expense and meat safety. Disease resistance in chickens can be improved through genetic selection for immunocompetence. Heterophils are found in the bone marrow, blood and connective tissue in poultry, and are the main leukocyte components in poultry. The function of heterophils is similar to that of mammalian neutrophils, which phagocytose and kill invading pathogens and protect the body from invasion[1]. Heterophils play a crucial role in killing foreign pathogenic microorganisms when the specific immune system of the body is not mature. H/L, the ratio of heterophil to lymphocyte, can be used to study the corresponding immune

effects directly at the cellular level. The heterophil/lymphocyte ratio (H/L) in chicken peripheral blood has been widely accepted as a reliable and accurate physiological indicator of the chicken stress response[2]. With high or low temperatures, bacterial infection and other stress reactions, the number of heterophils increases[3, 4], H/L can be used as a blood index reflecting the disease resistance of the body[5].

Genome-wide association study(GWAS) can provide preliminary results in genetic research on complex traits in animals. Many single nucleotide polymorphism (SNP) loci and qualitative trait loci(QTL) with strong statistical significance have been discovered[6-9]. GWAS has become a powerful tool for studying important economic traits of livestock and poultry and has promoted research on marker-assisted breeding. However, GWAS cannot be used to identify the causative variation directly association with the traits. The correlation signals that a linked regions. However, pinpointing the causative variation is more difficult. To reveal the mechanism through which a mutation site affects a phenotype and to perform subsequent functional research, GWAS joint analysis on multiple genomic levels can be used and biological pathway analysis can be applied to detect the superposition of multiple minor genes by examining genes involved in the same biological pathway, thus enabling deeper mining of GWAS data[10-13]. With the development of genome sequencing technology and the continuous improvement of statistical methods, GWAS is expected to be more efficiently applied to gene identification for important traits in livestock and poultry and to play an increasingly important role in animal breeding.

Our goal was to identify the candidate genes associated with H/L disease-resistance traits, with the goal of mitigating damage, and to conduct a pathway analysis to complement previously obtained GWAS results for phenotypes associated with H/L.

Materials And Methods

Ethics statement

The work was approved by the Animal Management Committee of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (IAS-CAAS, Beijing, China). Ethical approval regarding animal survival was granted by the animal ethics committee of IAS-CAAS (approval number: IASCAAS-AE20140615).

Phenotypes and breeding value estimation

In chickens at 42 days of age, 10 μ L of blood was freshly obtained from the wing. Slides were air-dried and dyed with May–Grunwald–Giemsa stain, and the H/L ratio were calculated. The basic data have been used in recent GWAS analysis, and details regarding the genotyping and phenotyping have been reported[14]. In this study, a conventional pedigree-based best linear unbiased prediction model (BLUP) was used to predict breeding values for each trait.

The pedigree-based BLUP model[15] is

$$y = Xb + Za + e$$

where y is the vector of the phenotypic records of the trait (H/L), b is the vector of the fixed effect (Batch and Sex), X is the incidence matrix linking b to y , a is a vector of additive breeding values to be estimated, Z is the incidence matrices linking a to y and e is the vector of residuals. We assumed that $\text{var}(a) = A\sigma_a^2$, where A is the pedigree-based genetic relationship matrix. The prediction of breeding values (ebv) was performed the Asreml package[16]. The ebv calculated in this part is used for the next GWAS analysis

Genotyping and quality control

Briefly, 1,317 birds were genotyped with Axiom® SNP arrays[17] with the Affymetrix GeneTitan® system according to the procedure described in the Axiom 2.0 Target Prep 384 Samples Protocol (Affymetrix; Santa Clara, CA, USA) (<https://www.thermofisher.com/>). Stringent quality control (QC) procedures were performed for the genotype data by using PLINK[18] version 1.09(<http://zzz.bwh.harvard.edu/plink/>), After QC[call rate >95%, minor allele frequency >0.05, and extreme deviation from Hardy-Weinberg proportions($P > 1E-6$)], 1,212 animals and 36,800 SNPs, located on 28 autosomes and in the Z-chromosome, were retained. Slight differences in the number of individuals and SNPs across the recent GWAS analyses are attributed to phenotypic editing.

Single SNP-based association analysis

The basic MLM model was used for association analysis in GAPIT[19]. The P-value was corrected with a strict Bonferroni adjustment based on LD pruning[20]. The sums of the independent LD blocks plus singleton markers were used to define the number of independent statistical comparisons [21]. Finally, 8,068 independent SNPs were used to determine the P-value thresholds, including 5% genome-wide significance ($6.80E-6, 0.05/8,068$), and suggestive association ($1.24E-4, 1/8,068$). The Quantile-quantile (Q-Q) plot and Manhattan plots of GWAS for H/L were produced with the “qqman” packages[22] in R. The most significant SNP genotypes were added to univariate models to analyze the phenotypic differences among individuals with different SNP genotypes.

Pathway-based association analysis

The pathway-based analysis workflow is represented in Figure 1. In brief, for each trait, nominal P-values < 0.05 from the GWAS analyses were used to identify significant SNPs. With VEP software (<http://asia.ensembl.org/Tools/VEP>), SNPs were assigned to genes if they were within the genomic sequence of a gene or within a flanking region of 5 kb up- or downstream of a gene, to include SNPs located in regulatory regions. The Ensembl GRCg6a database was used as a reference (http://asia.ensembl.org/Gallus_gallus/Info/Index). The background SNPs represent all SNPs tested in the GWAS analyses, whereas the background genes were the genes associated with those SNPs. For assignment of the genes to functional categories, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway[23] databases were used. Pathway enrichment analysis was performed with the OmicShare tools, a free online platform for data analysis (www.omicshare.com/tools). Significantly enriched pathways in significant genes to compared with the background genes were defined with a hypergeometric test.

Proportion of phenotypic variance explained by SNPs (PVE)

PVE estimation is equivalent to calculating the heritability of SNPs,

$$h_{snp}^2 = \sigma_g^2 / \sigma_p^2$$

The GBLUP model was used to estimate the phenotypic variance explained by SNPs of all genotyped individuals[24]:

$$y = Xb + Zg + e$$

where the definitions of y , X , Z and e are the same as those in the BLUP model, g is the vector of genomic breeding values to be estimated, following a normal distribution of $N(0, G\sigma_g^2)$, in which σ_g^2 is the variance of additive genetic effect, and G is the marker-based genomic relationship matrix. Z is the incidence matrix linking g to y , and e is the vector of residuals. e is the vector of random errors, following a normal distribution of $N(0, I\sigma_e^2)$, in which σ_e^2 is the residual variance. The estimation of variance components was performed with the Asreml package[16].

Results

Single SNP-based association study

In this study, 1,317 birds were genotyped with Affymetrix Chicken 55K genotyping arrays. According to the QC criteria, 1,212 animals and 36,750 SNPs were used in subsequent analyses. A total of 36,750 SNP markers were obtained with principal component analysis (PCA) by using the first three principal components (Figure 2). P-value correction was first performed with PLINKv1.9 software[19] for block analysis. Bonferroni correction was used to correct the P-values, and 8,068 independent linkage disequilibrium (LD) blocks were obtained.

GAPIT was used for traditional SNP-based association analysis. In total, one associated SNP (5% genome-wide significance (6.80E-6, 0.05/8068)) on GGA1 (chicken chromosome 1) and seven SNPs with a suggestive association (1.24E-4, 1/8068) on GGAs 1, 7, 13 and 19 were detected. The detailed information on the SNPs reaching 5% genome-wide significance and the seven SNPs with a suggestive association is shown in Table 1. A novel gene in the 100 Kb region of chromosome 1 plus or minus. HNRNPA3 (Heterogeneous Nuclear Ribonucleoprotein A3) is a protein coding gene that plays a role in the cytoplasmic trafficking of RNA. PLEKHM3 (Pleckstrin Homology Domain Containing M3) may act as a scaffold protein for AKT1 during muscle differentiation. C1QBP (Complement C1q Binding Protein) was believed to be a multifunctional and multicompartamental protein involved in inflammation and infection processes, and its signaling has been implicated in inhibition of the innate immune response via the PI3K-AKT/PKB pathway. TENM2 (Teneurin Transmembrane Protein 2) is a protein coding gene that mediates axon guidance and heterophilic cell-cell adhesion. ZNF385B (Zinc Finger Protein 385B) may

play a role in p53/TP53-mediated apoptosis. The Q-Q plot is presented in Figure 3. The global view of P-values (in terms of $-\log_{10}$ (P-value)) for all SNPs is represented by a Manhattan plot in Figure 4. The raw results of all SNPs are described in Additional file 1: Table S1.

Table 1. SNPs with genome-wide significance for H/L.

Traits	SNP	chromosome	Position(BP)	P-value	Nearest Gene
H/L_ebv	AX_172586886	1	1.89E+08	4.23E-06	ENSGALG00000041225 (U52kb)
H/L_ebv	AX_76986680	7	15894092	2.20E-05	HNRNPA3 (within)
H/L_ebv	AX_172565168	7	12098710	2.53E-05	PLEKHM3 (within)
H/L_ebv	AX_75926692	19	3405344	3.18E-05	C1QBP (within)
H/L_ebv	AX_75383308	1	1.89E+08	4.83E-05	ENSGALG00000041225 (U21kb)
H/L_ebv	AX_172582339	13	5407991	5.28E-05	TENM2 (D102Kb)
H/L_ebv	AX_75383086	1	1.89E+08	8.56E-05	ENSGALG00000041225 (within)
H/L_ebv	AX_76984294	7	14956470	9.37E-05	ZNF385B (U24Kb)

To effectively identify SNPs, We performed additional analyses, and the SNPs significantly associated with H/L and H/L ebv were fitted into the model to examine these associations. We suggest that the five identified loci provide the most reliable signal. The substitution of variant wild-type for mutant SNPs led to a significant decrease in H/L and H/L ebv. The results for all SNPs are described in Additional file 2: Figure S1.

Pathway-based association study

The same genotyping dataset was used in this analysis. On the basis of the QC and SNP selection criteria (Methods), 335 pathways were selected from the KEGG dataset, covering 13,068 genes. In brief, nominal P-values < 0.05 from the GWAS analyses were used to identify significant SNPs. In total, 1,839 SNP (out of the 36750 tested) were located in annotated genes or in the 10 kb window (up-stream or down-stream of a gene). The H/L trait was analyzed, and we identified nine associated pathways ($P < 0.05$): small cell lung cancer, proteoglycans in cancer, calcium signaling pathway, microRNAs in cancer, sulfur relay system, pathways in cancer, gastric acid secretion, purine metabolism and salivary secretion. The Q-Q plot is shown in Figure 5. The detailed information for each pathway is provided in Additional file: Table S2 and Table S3. Furthermore, we prepared plots of the number of genes for each pathway class to investigate the gene function. The results of gene classification confirmed the accuracy of the GWAS analysis (Additional file 4: Figure S2).

Estimation of phenotypic variance explained by SNPs from various datasets

To determine the contribution of SNPs to the phenotypic variation in H/L, we estimated the phenotypic variance explained by SNPs from various datasets (Table S4, Figure 6). We analyzed all SNPs after QC

(QC_SNP, 36,750 SNPs) and found that the phenotypic variation that they were able to explain was 12.4%. The phenotypic variation explained by GWAS-based loci (5 SNPs) and pathway-based loci (48 SNPs) was 4.2% and 9.1%, respectively. The phenotypic variation in H/L that could be explained by SNPs (53 SNPs, GWAS-based and pathway-based) identified jointly by the two analysis was as high as 9.7%.

Discussion

Key genes associated with the H/L ratio

Heterophils are a special type of white blood cells in poultry, which are responsible for immune defense and immune regulation in the body's innate immune system. Similarly to mammalian neutrophils, heterophils mainly kill pathogens through trapping, phagocytosis, oxidative bursts, cytokine production and other processes, thus protecting poultry from invasion[1]. The process through which phagocytic leukocytes produce reactive oxygen species (ROS) under external stimuli (such as microorganisms and microorganism-related molecules) is called oxidation or the respiratory burst. When the body is infected by pathogenic microorganisms and produces an acute inflammatory reaction, heterophils migrate to the inflammatory site. C1QBP mediates many biological responses in the immune system, including inflammation, infection and immune regulation[25]. Mechanistically, C1QBP is a direct target gene of ZNF32; it inactivates the p38 mitogen-activated protein kinase (MAPK) pathway, thereby exerting the protective effects of ZNF32 against oxidative stress-induced apoptosis. C1QBP is essential for sustaining mitochondrial membrane potential, and it enhances cellular resistance to oxidative stress[26]. C1QBP promotes the release of cytokines such as TNF- α , IL-1 β , IL-6, IL-8 and MCP-1 from mononuclear cells[27, 28]. The C1QBP-enhanced phagocytic process occurs through the activity of phosphoinositide 3-kinase (PI3K)[29, 30]. C1QBP also plays an essential role in CSF-1 induced migration of macrophages[31].

ENSGALG00000041225 is a novel gene, encoding SH3 domain-containing protein. The SRC Homology 3 Domain (or SH3 domain) is a small protein domain of approximately 60 amino acid residues. This domain has been identified in several protein families such as PI3K, Ras GTPase-activating protein, CDC24 and CDC25. SH3 domains are found in proteins involved in signaling pathways regulating the cytoskeleton, such as the Ras protein, the Src kinase and many others. Because PI3K and Src kinase are key enzymes in heterophils, this gene may be associated with heterophil function[32].

Signaling pathways associated with the H/L ratio

The regulation of the H/L ratio may be a result of complex interactions among multiple pathways. According to the pathway-based analysis, several well-known pathways associated with immune response were found to be enriched (e.g., small cell lung cancer, proteoglycans in cancer, microRNAs in cancer, pathways in cancer, etc.). In addition, other signaling pathways (calcium signaling pathway, sulfur relay system, purine metabolism and salivary secretion) were enriched significantly ($P < 0.05$).

The sulfur relay system pathway, regulated by MPST and URM1, was significantly associated with H/L. According to previous research, MPST catalyzes the transfer of a sulfur ion from 3-mercaptopyruvate to

cyanide or other thiol compounds. MPST is required for thiosulfate biosynthesis and it contributes to the catabolism of cysteine and is an important producer of hydrogen sulfide(H_2S)[33, 34]. Treatment with H_2S increase the proliferation and viability of LPS-treated L6 cells. H_2S decrease ROS- induced apoptosis through the MAPK signaling pathway in LPS-treated L6 cells. H_2S ameliorates LPS-induced inflammation through the ROS/MAPK and TLR4/NF- κ B signaling pathways[35]. Both endogenous and exogenous H_2S inhibits cell proliferation and cytokine release, and inhibits ERK-1/2 and MAPKs, thus resulting in a clear anti-inflammatory response[36]. H_2S influences multiple biological functions of cells through inhibiting the PI3K/Akt/mTOR signaling pathway and consequently cell migration, proliferation and cell division[37].

Contribution of SNPs to the variation in H/L

The dissection of the H/L effect was based on the theory of quantitative polygenes micro-effects. H/L is a complex trait affected by multiple genetic loci, but SNPs significantly associated with H/L, according to GWAS analysis results, can usually explain only part of phenotypic variation. This partial explanation is primarily because of interactions among genes or between genes and the environment, and the phenotypic variation that can be explained by SNPs is usually much lower than that of heritability in a narrow sense (e.g., the ratio of additive genetic variance to phenotypic variance). In contrast, non-additive effects do not contribute to narrow-sense heritability, mainly because some causative variation sites affecting phenotypes do not reach the threshold of significance, or the causative variation is not in LD with SNPs that have been genotyped[38]. An other reason is that GWAS considers SNPs above a significance threshold, but a lack of SNPs that reach the threshold does not mean that there is no additive effect. We found that all SNPs after QC explained 12.4% of the phenotypic variation in H/L, and the phenotypic variation explained by phenotypic SNPs was between 4.2% and 9.7%. In the H/L GWAS, combined with the pathway correlation analysis, 53 SNPs associated with H/L explained as much as 9.7% of the phenotypic variation in H/L, a value close to the phenotypic variation in H/L that all SNPs were able to explain. This results also precisely indicates that SNPs significantly associated with H/L play a major role in the variation in H/L in all SNPs. Simultaneously, correlation analysis based on pathways can meaningfully supplement the results of GWAS analysis.

Conclusion

This study comprised an SNP- and pathway-based analysis to identify genes associated with chicken H/L traits, revealing genes associated with H/L. The gene C1QBP, identified to be associated with H/L content, may be an important candidate gene involved in the regulation of H/L. We also provided five SNPs that may be used as markers in breeding programs. To our knowledge, this is the first pathway-based association analysis of H/L traits and GWAS analysis related to H/L ebv. In GWASs and animal breeding, studies have mainly focused on simple associations of SNPs with traits. The pathway-based analysis provided new findings relative to previous analyses, and confirmed that complex traits are affected by the cumulative effects of several genes clustered in specific biological pathways. The

pathways found to be associated with traits may be useful in further studies involving gene mapping and developing of breeding programs using SNP markers. Our findings provide new insights into the genetic mechanism underlying H/L.

Abbreviations

H/L Heterophil/lymphocyte ratio

GWAS Genome-wide association study

QTL Qualitative trait loci

BLUP Best linear unbiased prediction model

QC Quality control

KEGG Kyoto Encyclopedia of Genes and Genomes

PVE Proportion of phenotypic variance explained by SNPs

PCA Principal component analysis

LD Linkage disequilibrium

Declarations

Acknowledgments

The authors would like to thank all staffs at the the Xinguang Agriculture and Animal Husbandry Co., Ltd in Foshan for animal care and sample collection. We thank International Science Editing (<http://www.internationalscienceediting.com>) for editing this manuscript.

Funding

This research was funded by the China Agriculture Research System (CARS42), the National High-Tech R&D Program (2015BAD03B03) and the Agricultural Science and Technology Innovation Program (ASTIPIAS04) of the Chinese Academy of Agricultural Sciences.

Availability of data and materials

The data and computing programs used in this manuscript are available from the corresponding author on request.

Author information

Affiliations

State Key Laboratory of Animal Nutrition, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing, 100193, China

Jie Wang, Bo Zhu, Jie Wen, Qinghe Li and Guiping Zhao

Contributions

QL and GZ conceived and designed the study. JW performed statistical analyses and wrote the paper. BZ participated in data analyses. JW and BZ participated in the design of the study and contributed to acquisition of data. All authors read and approved the final manuscript.

Corresponding authors

Correspondence to Guiping Zhao or Qinghe Li.

Ethics declarations

Ethics approval and consent to participate

Animal experiments were approved by the Science Research Department of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (CAAS) (Beijing, China). There was no use of human participants, data or tissues.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Rights and permissions

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

References

1. Redmond SB, Chuammitri P, Andreasen CB, Palić D, Lamont SJ: Genetic control of chicken heterophil function in advanced intercross lines: associations with novel and with known Salmonella resistance loci and a likely mechanism for cell death in extracellular trap production. *immunogenetics* 2011, 63(7):449-458. DOI: 10.1007/S00251-011-0523-Y.
2. Al-Murrani WK, Kassab A, Al-Sam HZ, Al-Athari AMK: Heterophil/lymphocyte ratio as a selection criterion for heat resistance in domestic fowls. *british poultry science* 1997, 38(2):159-163. DOI: 10.1080/00071669708417962.
3. McFarlane JM, Curtis SE: Multiple Concurrent Stressors in Chicks.3. Effects on Plasma Corticosterone and the Heterophil:Lymphocyte Ratio. *poultry science* 1989, 68(4):522-527. DOI: 10.3382/PS.0680522.
4. Gross WB: Factors affecting chicken thrombocyte morphology and the relationship with heterophil: Lymphocyte ratios. *british poultry science* 1989, 30(4):919-925. DOI: 10.1080/00071668908417218.
5. Masello JF, Choconi RG, Helmer M, Kremberg T, Lubjuhn T, Quillfeldt P: Do leucocytes reflect condition in nestling burrowing parrots *Cyanoliseus patagonus* in the wild. *comparative biochemistry and physiology a molecular & integrative physiology* 2009, 152(2):176-181. DOI: 10.1016/J.CBPA.2008.09.018.
6. Zhou Z, Li M, Cheng H, Fan W, Yuan Z, Gao Q, Xu Y, Guo Z, Zhang Y, Hu J *et al*: An intercross population study reveals genes associated with body size and plumage color in ducks. *nature communications* 2018, 9(1). DOI: 10.1038/S41467-018-04868-4.
7. Liu JJ, Liang AX, Campanile G, Plastow G, Zhang C, Wang Z, Salzano A, Gasparini B, Cassandro M, Yang LG: Genome-wide association studies to identify quantitative trait loci affecting milk production traits in water buffalo. *journal of dairy science* 2018, 101(1):433-444. DOI: 10.3168/JDS.2017-13246.
8. Eusebi PG, González-Prendes R, Quintanilla R, Tibau J, Cardoso TF, Clop A, Amills M: A genome-wide association analysis for carcass traits in a commercial Duroc pig population. *animal genetics* 2017, 48(4):466-469. DOI: 10.1111/AGE.12545.
9. Azmal SA, Bhuiyan AA, Omar AI, Ma S, Sun C, Han Z, Zhang M, Zhao S, Li S: Novel Polymorphisms in RAPGEF6 Gene Associated with Egg-Laying Rate in Chinese Jing Hong Chicken using Genome-Wide SNP Scan. *genes* 2019, 10(5). DOI: 10.3390/GENES10050384.
10. Walsh N, Zhang H, Hyland PL, Yang Q, Mocci E, Zhang M, Childs EJ, Collins I, Wang Z, Arslan AA *et al*: Agnostic Pathway/Gene Set Analysis of Genome-Wide Association Data Identifies Associations for Pancreatic Cancer. *journal of the national cancer institute* 2019, 111(6):557-567. DOI: 10.1093/JNCI/DJY155.
11. White MJ, Yaspan BL, Veatch OJ, Goddard P, Risse-Adams OS, Contreras MG: Strategies for Pathway Analysis Using GWAS and WGS Data. *current protocols in human genetics* 2019, 100(1). DOI: 10.1002/CPHG.79.
12. Kwak IY, Pan W: Gene- and pathway-based association tests for multiple traits with GWAS summary statistics. *bioinformatics* 2017, 33(1):64-71. DOI: 10.1093/BIOINFORMATICS/BTW577.

13. Ostrom QT, Coleman W, Huang W, Rubin JB, Lathia JD, Berens ME, Speyer G, Liao P, Wrensch MR, Eckel-Passow JE *et al*: Sex-specific gene and pathway modeling of inherited glioma risk. *neuro oncology* 2019, 21(1):71-82. DOI: 10.1093/NEUONC/NOY135.
14. Zhu B, Li Q, Liu R, Zheng M, Wen J, Zhao G: Genome-Wide Association Study of H/L Traits in Chicken. *animal* 2019, 9(5). DOI: 10.3390/ANI9050260.
15. Cr H: Best linear unbiased estimation and prediction under a selection model. *biometrics* 1975, 31(2):423-447. DOI: 10.2307/2529430.
16. Burgueño J, Cadena A, Crossa J, Banziger M, Gilmour AR, Cullis B: User's guide for spatial analysis of field variety trials using ASREML. In.; 2000.
17. Liu R, Xing S, Wang J, Zheng M, Cui H, Crooijmans RPMA, Li Q, Zhao G, Wen J: A new chicken 55K SNP genotyping array. *bmc genomics* 2019, 20(1). DOI: 10.1186/S12864-019-5736-8.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, Bakker PIWd, Daly MJ *et al*: PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *american journal of human genetics* 2007, 81(3):559-575. DOI: 10.1086/519795.
19. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z: GAPIT: Genome Association and Prediction Integrated Tool. *bioinformatics* 2012, 28(18):2397-2399. DOI: 10.1093/BIOINFORMATICS/BTS444.
20. Johnson RC, Nelson GW, Troyer JL, Lautenberger JA, Kessing BD, Winkler CA, O'Brien SJ: Accounting for multiple comparisons in a genome-wide association study (GWAS). *bmc genomics* 2010, 11(1):724-724. DOI: 10.1186/1471-2164-11-724.
21. Bai H, Sun Y, Liu N, Xue F, Li Y, Xu S, Ye J, Zhang L, Chen Y, Chen J: Single SNP- and pathway-based genome-wide association studies for beak deformity in chickens using high-density 600K SNP arrays. *bmc genomics* 2018, 19(1):501-501. DOI: 10.1186/S12864-018-4882-8.
22. Turner SD: qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. In., vol. 3; 2018.
23. Kanehisa M, Goto S: KEGG: Kyoto Encyclopedia of Genes and Genomes. *nucleic acids research* 1999, 28(1):27-30. DOI: 10.1093/NAR/27.1.29.
24. VanRaden PM: Efficient methods to compute genomic predictions. *journal of dairy science* 2008, 91(11):4414-4423. DOI: 10.3168/JDS.2007-0980.
25. Peerschke EIB, Minta JO, Zhou SZ, Bini A, Gotlieb A, Colman RW, Ghebrehiwet B: Expression of gC1q-R/p33 and its major ligands in human atherosclerotic lesions. *molecular immunology* 2004, 41(8):759-766. DOI: 10.1016/J.MOLIMM.2004.04.020.
26. Li K, Gao B, Li J, Chen H, Li Y, Wei Y, Gong D, Gao J, Zhang J, Tan W *et al*: ZNF32 protects against oxidative stress-induced apoptosis by modulating C1QBP transcription. *oncotarget* 2015, 6(35):38107-38126. DOI: 10.18632/ONCOTARGET.5646.
27. Ghebrehiwet B, Jesty J, Xu S, Vinayagasundaram R, Vinayagasundaram U, Ji Y, Valentino A, Hosszu KK, Mathew S, Joseph K *et al*: Structure–Function Studies Using Deletion Mutants Identify Domains

- of gC1qR/p33 as Potential Therapeutic Targets for Vascular Permeability and Inflammation. *frontiers in immunology* 2011, 2. DOI: 10.3389/FIMMU.2011.00058.
28. Khan MM, Bradford HN, Isordia-Salas I, Liu Y, Wu Y, Espinola RG, Ghebrehiwet B, Colman RW: High-Molecular-Weight Kininogen Fragments Stimulate the Secretion of Cytokines and Chemokines Through uPAR, Mac-1, and gC1qR in Monocytes. *arteriosclerosis thrombosis and vascular biology* 2006, 26(10):2260-2266. DOI: 10.1161/01.ATV.0000240290.70852.C0.
29. Shen Y, Naujokas M, Park M, Ireton K: InIB-Dependent Internalization of Listeria Is Mediated by the Met Receptor Tyrosine Kinase. *cell* 2000, 103(3):501-510. DOI: 10.1016/S0092-8674(00)00141-0.
30. Bierne H, Gouin E, Roux P, Caroni P, Yin HL, Cossart P: A role for cofilin and LIM kinase in Listeria-induced phagocytosis. *journal of cell biology* 2001, 155(1):101-112. DOI: 10.1083/JCB.200104037.
31. Wang Y, Su J, Yuan B, Fu D, Niu Y, Yue D: The role of C1QBP in CSF-1-dependent PKC ζ activation and macrophage migration. *experimental cell research* 2018, 362(1):11-16. DOI: 10.1016/J.YEXCR.2017.09.038.
32. Berton G, Mócsai A, Lowell CA: Src and Syk kinases: key regulators of phagocytic cell activation. *trends in immunology* 2005, 26(4):208-214. DOI: 10.1016/J.IT.2005.02.002.
33. Kuo MM, Kim DH, Jandu S, Bergman Y, Tan S, Wang H, Pandey DR, Abraham TP, Shoukas AA, Berkowitz DE *et al*: MPST but not CSE is the primary regulator of hydrogen sulfide production and function in the coronary artery. *american journal of physiology heart and circulatory physiology* 2016, 310(1). DOI: 10.1152/AJPHEART.00574.2014.
34. Baskin SI, Wing DA, Kirby SD: Cyanide conversion to thiocyanate by 3-mercaptopyruvate sulfurtransferase (MPST). In.; 1990.
35. Zhang G-Y, Lu D, Duan S-F, Gao Y-R, Liu S-Y, Hong Y, Dong P-Z, Chen Y-G, Li T, Wang D-Y *et al*: Hydrogen Sulfide Alleviates Lipopolysaccharide-Induced Diaphragm Dysfunction in Rats by Reducing Apoptosis and Inflammation through ROS/MAPK and TLR4/NF- κ B Signaling Pathways. *oxidative medicine and cellular longevity* 2018, 2018. DOI: 10.1155/2018/9647809.
36. Perry MM, Tildy B, Papi A, Casolari P, Caramori G, Rempel KL, Halayko AJ, Adcock I, Chung KF: The anti-proliferative and anti-inflammatory response of COPD airway smooth muscle cells to hydrogen sulfide. *respiratory research* 2018, 19(1):85-85. DOI: 10.1186/S12931-018-0788-X.
37. Wang SS, Chen YH, Chen N, Wang LJ, Chen DX, Weng HL, Dooley S, Ding HG: Hydrogen sulfide promotes autophagy of hepatocellular carcinoma cells through the PI3K/Akt/mTOR signaling pathway. *cell death and disease* 2017, 8(3). DOI: 10.1038/CDDIS.2017.18.
38. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW *et al*: Common SNPs explain a large proportion of the heritability for human height. *nature genetics* 2010, 42(7):565-569. DOI: 10.1038/NG.608.

Figures

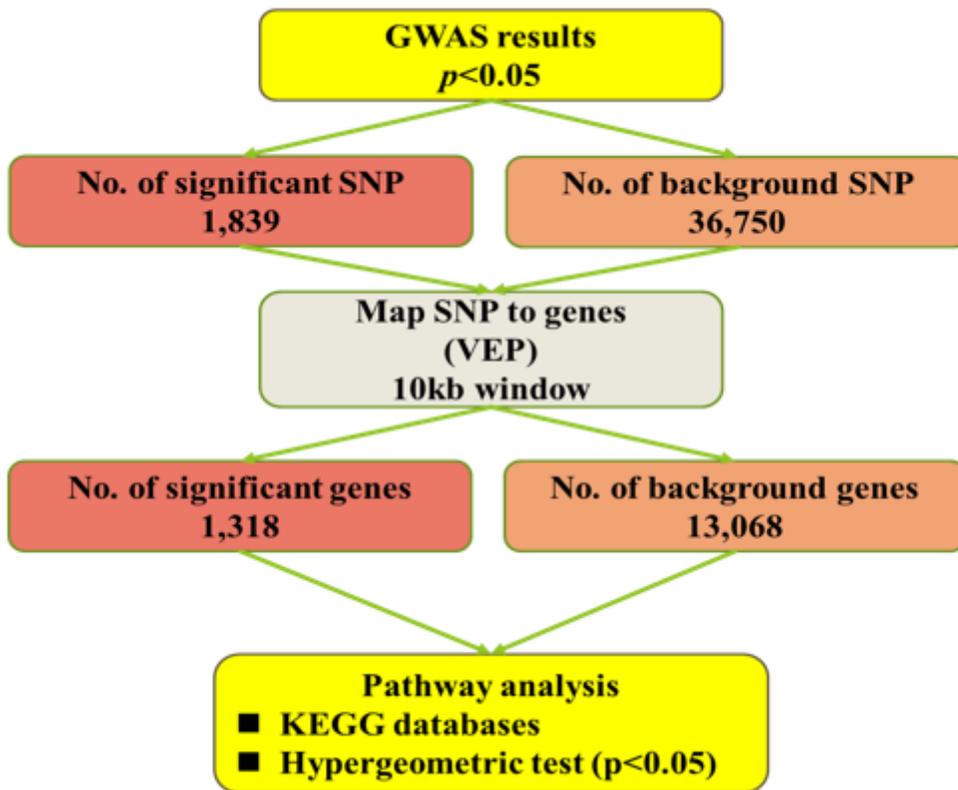


Figure 1

Flowchart for pathway-based analysis. GWAS= genome-wide association study; KEGG = Kyoto Encyclopedia of Genes and Genomes.

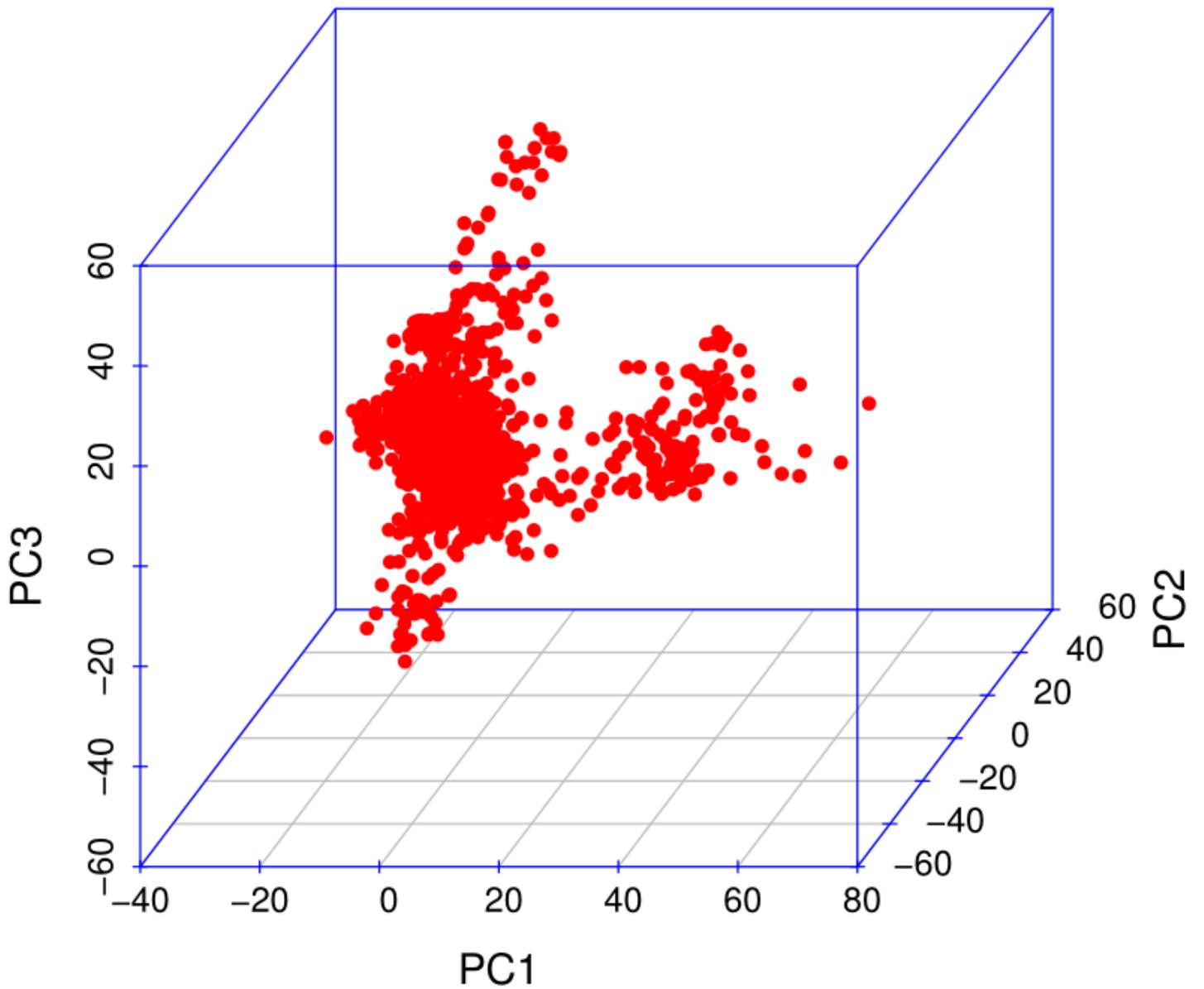


Figure 2

Population structure evaluated on the basis of the first three principal components.

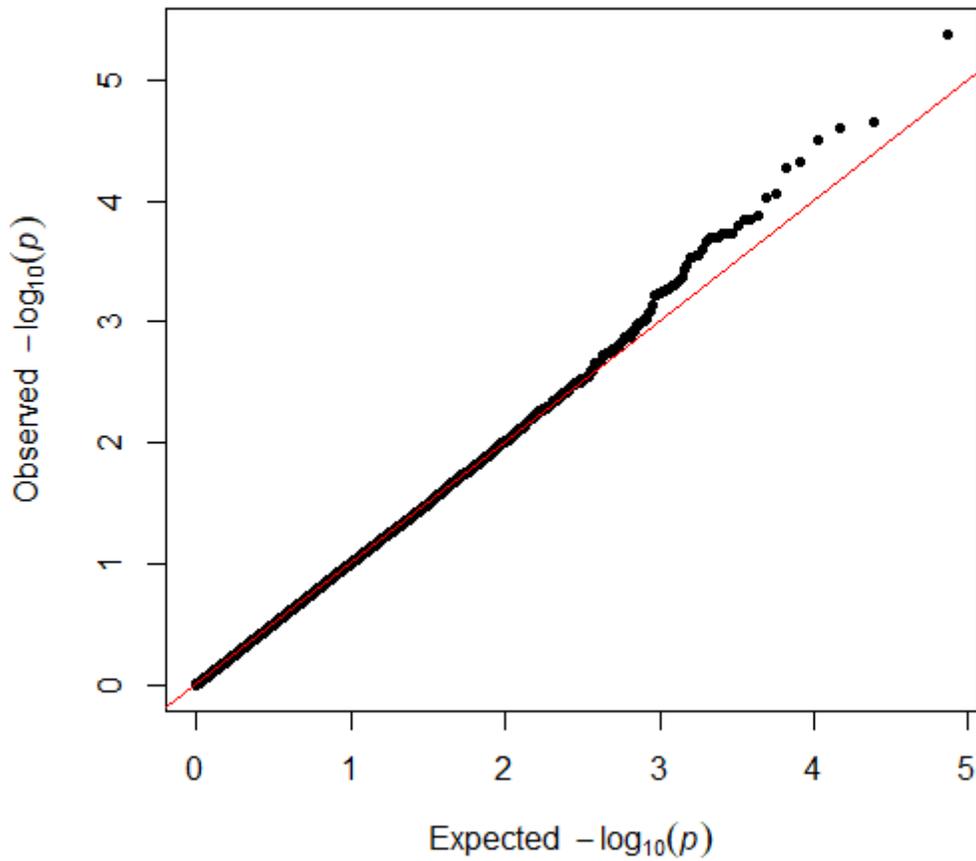


Figure 3

Q-Q plot of the GWAS results with GAPIT 3.0. The x-axis shows the expected P-values under the null hypothesis and the y-axis shows the observed P-values.

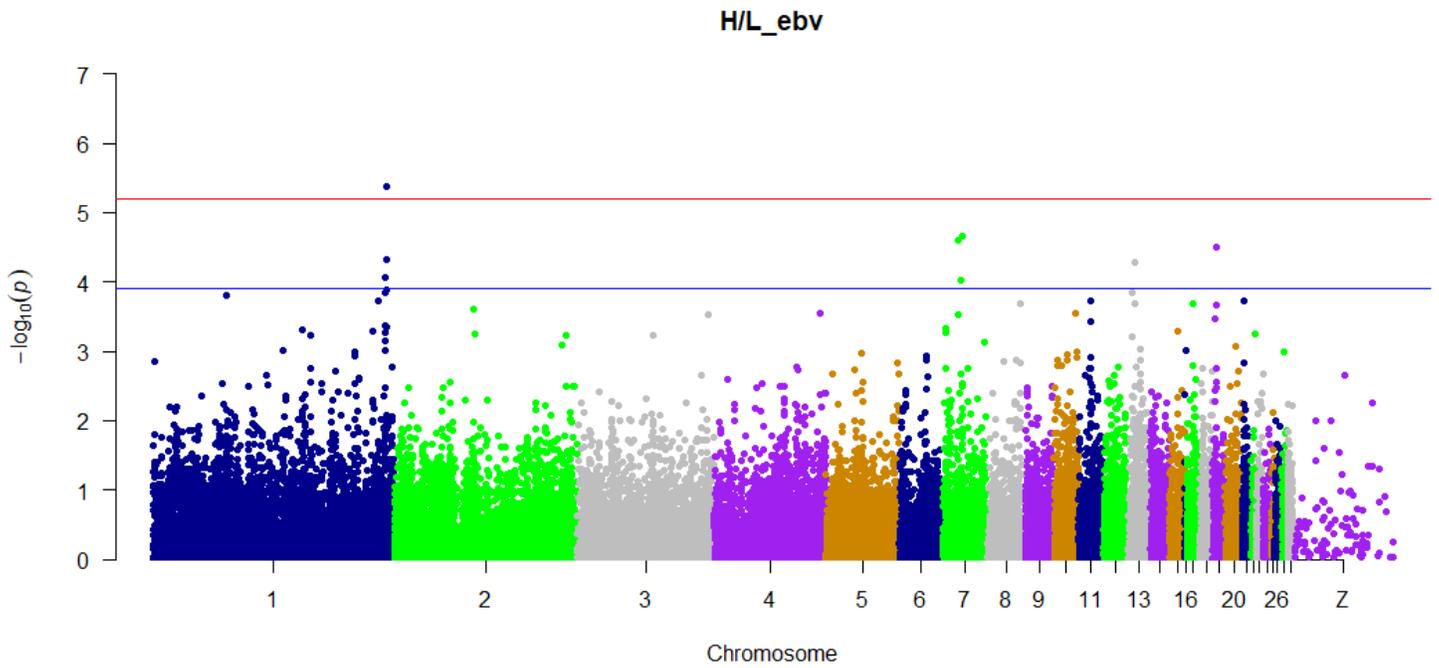


Figure 4

Manhattan plots showing association of all SNPs with the H/L ratio trait, by using GAPIT 3.0. SNPs are plotted on the x-axis according to their positions on each chromosome, and their association with H/L is shown on the y-axis (as $-\log_{10}(P\text{-value})$). The blue dashed line indicates suggestive genome-wide significance ($P\text{-value} = 1.24E-4$), and the red line shows genome-wide 5% significance with a $P\text{-value}$ threshold of $6.80E-6$.

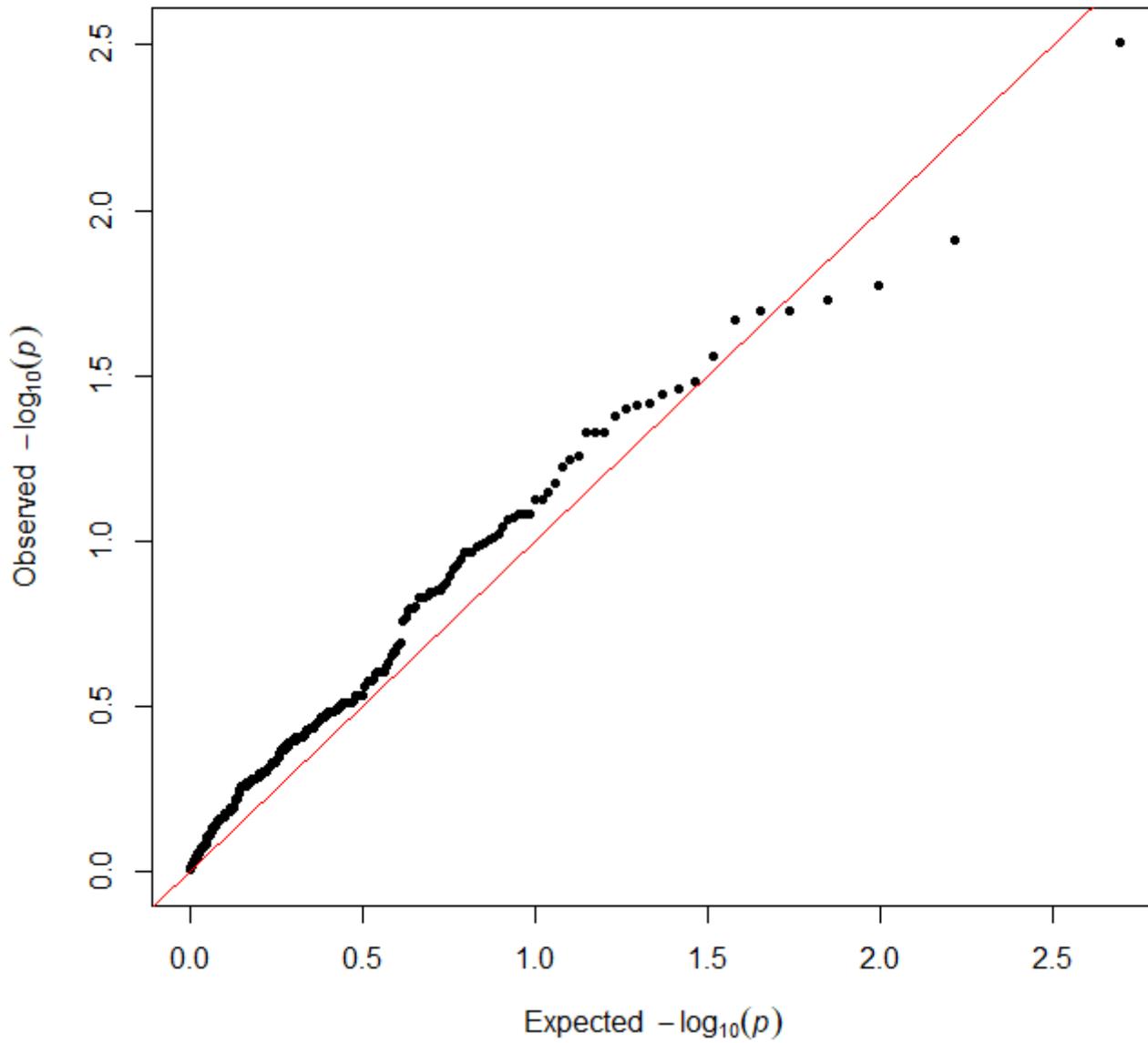


Figure 5

Q-Q plot for the SNPs used in the pathway-based analysis. The x-axis shows the expected P-values under the null hypothesis and the y-axis shows the observed P-values.

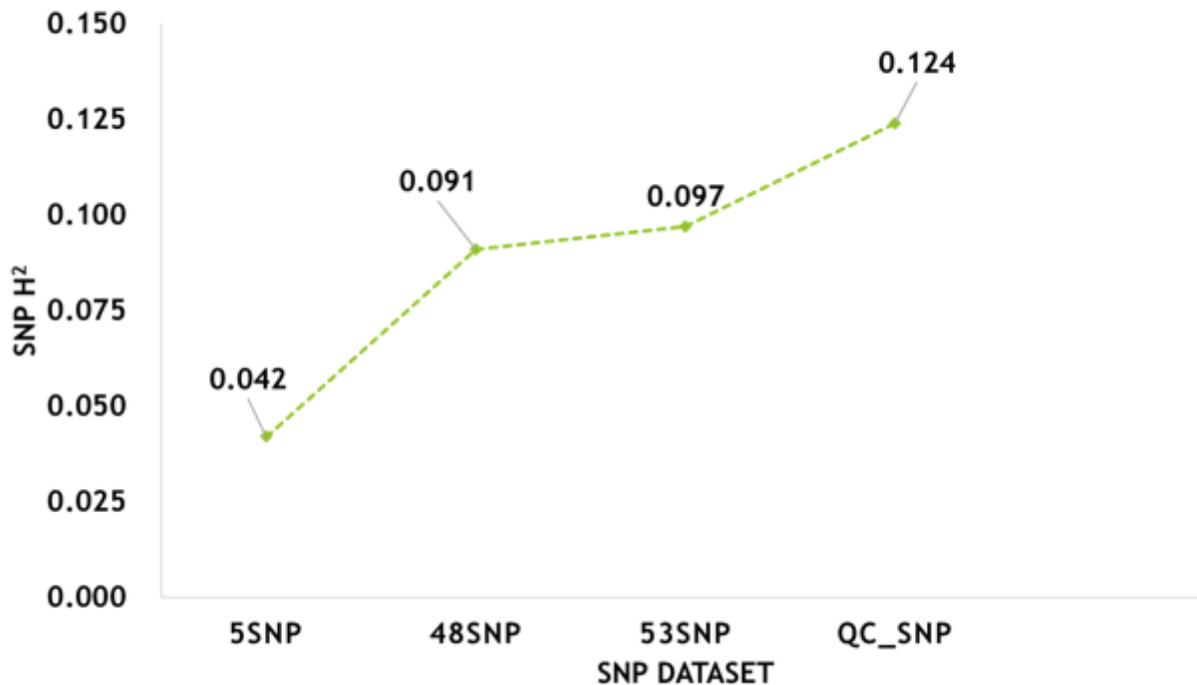


Figure 6

The contribution of SNPs from different datasets to the variation in H/L

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FiguresS1..docx](#)
- [FiguresS1..docx](#)
- [TableS1RawResultsofAlltheSNPs.xlsx](#)
- [TableS1RawResultsofAlltheSNPs.xlsx](#)
- [FiguresS2.docx](#)
- [FiguresS2.docx](#)
- [TableS2PathwayEnrichment.xlsx](#)
- [TableS2PathwayEnrichment.xlsx](#)
- [TableS3PathwayDetail.xlsx](#)
- [TableS3PathwayDetail.xlsx](#)
- [TableS4Estimationofphenotypicvariance.xlsx](#)
- [TableS4Estimationofphenotypicvariance.xlsx](#)