Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics

Loci associated with bone strength in laying hens

Biaty Raymond

Examensarbete / Swedish University of Agricultural Sciences
Department of Animal Breeding and Genetics
488
Uppsala 2016

Master's Thesis, 30 hp

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics

Loci associated with bone strength in laying hens

Biaty Raymond

Supervisors:

D.J. De-Koning, SLU, Department of Animal Breeding and Genetics
Anna Maria Johansson, SLU, Department of Animal Breeding and Genetics

Examiner:

Erling Strandberg, SLU, Department of Animal Breeding and Genetics

Credits: 30 hp
Course title: Degree project in Animal Science
Course code: EX0556
Level: Advanced, A2E

Place of publication: Uppsala
Year of publication: 2016
Name of series: Examensarbete / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, 488
On-line publicering: http://epsilon.slu.se

Key words: Osteoporosis, GWAS, Fracture, Genes, QTLs

ABSTRACT

One of the growing welfare concern in the layer industry is the high incidences of bone fracture. This is thought to result from reduction in bone strength due to osteoporosis which is exacerbated by environmental stresses and mineral deficiencies. Despite these factors however, the primary cause of bone weakness and the resulting fractures is believed to be genetic predisposition. In this study, we performed a genome-wide association study to identify with high reliability the loci associated with bone strength in laying hens. Genotype information and phenotype data were obtained from 752 laying hens belonging to the same pure line population. These hens were genotyped for 580,961 single nucleotide polymorphisms (SNPs) with each of the SNPs associated with tibial breaking strength using the family-based score test for association (FASTA). A total of 52 SNPs across chromosomes 1, 3, 8 and 16 were significantly associated with tibial strength with the genome-wide significance threshold set as a corrected p.value of 10e-5. Based on the local linkage disequilibrium around the significant SNPs, 5 distinct and novel QTLs were identified on chromosomes 1 (2 QTLs), 3 (1 QTL), 8 (1QTL) and 16 (1 QTL). The strongest association was detected within the QTL region on chromosome 8 with the most significant SNP having a corrected p.value of 4e-7. A number of candidate genes were identified within the QTL regions, including the BRD2 gene which is required for normal bone physiology. Bone-related pathways involving some of the genes were also identified including the hedgehog signalling and Wnt signalling pathways. Our result supports previous studies, which suggested that bone strength is highly regulated by genetics. It is therefore possible to alleviate bone fracture in laying hens through genetic selection, and ultimately improve hen welfare.

TABLE OF CONTENTS

BACKGROUND

Bone weakness resulting from osteoporosis represents a considerable welfare and economic problem in the layer industry. It is a pathological condition caused by a progressive loss in the amount of mineralised structural bone during the laying period. As a result, the bones become weak, with increased fragility and susceptibility to fracture. This condition is widespread in all laying flocks. It was estimated that about 98% of processed hens have fractured bones, and that 30% of commercial egg laying hens experience at least one incidence of bone fracture during the laying period even prior to depopulation and processing (Gregory & Wilkins 1989). Bone fracture is considered a welfare problem because of the acute and chronic pains associated with broken bones and the skeletal deformities that often remain from improperly healed fractures (Webster 2004). Economically, bone fractures impacts production and income by its effect on the incidence of mortality thereby decreasing the egg production output of farmers. A study reported that one of the major cause of mortality in caged laying hens was bone fracture (Weber 2003). Moreover, spent layers are usually marketed as a cheap source of meat, but their market value is reduced when bone fragments resulting from fractured bones are lodged in the meat.

Causes of bone weakness in laying hens

Over the years, several studies have attempted to understand the causal reasons for bone fractures in laying hens, and one of the first reasons put forward was mineral depletion in the bones. Because of its high demand for eggshell formation for example, Calcium is often mobilised from structural bone especially when dietary supply is inadequate, thereby leaving the hens with weak bones characterised by osteoporosis (Riddell 1992). The confinement of birds in limited spaces such as the battery cage housing system limits the ability of the hens to exercise. This limitation results in osteoporosis due to disuse and hence the consequent high incidences of fracture under such conditions. It has been shown that birds with ability to exercise in an aviary environment have stronger bones and lower cases of bone fracture than those confined in cages (Fleming & McCormack 2006). The way and manner in which hens are handled especially during depopulation and processing also impacts the incidences of bone fractures. A study showed that about 16-25% of hens suffer broken bones during the process in which they are removed from cages, and about 30% of hens experience new fractures during loading and transportation to processing facilities (Gregory & Wilkins 1989; Gregory & Wilkins 1990; Gregory et al. 1994). Although all reasons mentioned above can

exacerbate bone weakness and fractures in laying flocks, the primary cause is believed to be genetic in nature (Fleming & McCormack 2006; Bishop et al. 2000).

In the modern day layers, selection pressure has been on higher egg production with lower body weights and feed intake. Given that bone strength was not accounted for in past selection programs, it is hypothesised that over several generations of genetic selection for traits such as high egg production, bone strength has been negatively affected resulting in birds with genetically weaker bones. If indeed this is the case, it is possible then to reverse the condition with genetic selection for bone strength in laying hens.

The first study to prove that bone strength could be increased through genetic selection in poultry was by Mandour et al. (1989). In their selection experiment, they showed that after three generation of selection for humeral strength in broilers, the selected line had higher humeral strength than the control line. In a more recent study, Bishop et al. (2000) reported the inheritance of bone characteristics linked to bone strength in laying hens. In the study, they found that about 40% of the variation in the bone strength phenotype (in this case, the total bone index made up of 4 underlying trait components which were: keel radiographic density, humeral strength, tibial strength and body weight) was explained by genetic differences between the hens. Two lines selected for low and high bone index clearly differed in bone strength characteristic after just five generation of selection, with significant reduction in the incidence of bone fracture in the line selected for high bone index.

These results clearly indicate that through genetic selection and in the long term, it is possible to alleviate the problem of osteoporosis in laying hens. But in the study of Bishop et al. (2000), bone characteristics were measured retrospectively i.e. after the hens were slaughtered at the end of their laying period. That means only the progeny could be selected based on the bone characteristics of their parents. This phenotyping procedure cannot be implemented in commercial breeding programs. Other non-invasive procedures exist for measuring bone characteristics in laying hens (Schreiweis et al. 2005; Schreiweis & Orban 2003), which could be used to select candidates on which phenotypes are measured. These procedures however are either too expensive, labour or time demanding. A viable alternative for genetic selection that does not require phenotyping and can be implemented on a commercial scale is marker assisted selection. If markers can be identified which are linked to bone strength, they can be incorporated into genomic selection programs and will even

give faster response to selection due to the higher accuracy of genomic selection compared to classical BLUP based on only parent information (Meuwissen 2007).

Quantitative trait loci (QTLs) for bone strength found so far

A few studies have reported the result of genome scan for QTL regions that are linked to bone strength in laying hens. Dunn & Fleming (2007) reported a significant QTL found on chromosome 1 linked to bone index and the underlying trait components of the index. In the study, they used an F2 cross between two lines divergently selected for bone index as described in Bishop et al. (2000). Melissa et al. (2005) reported multiple suggestive QTLs for bone strength, although no significant QTL for bone mineral density and content was found after adjusting for body weight and egg production. Rubin et al. (2007) also reported about four significant QTLs for femoral traits with one of the QTL directly associated with bone strength in laying hens. However, all these studies used a considerably few number of markers for detecting QTLs (ranging from 120 to 164 microsatellite markers). When only a few markers are used, the statistical power of detecting an association is considerably low (Spencer et al. 2009), which means that there are possibly important regions especially those with smaller effects that could not be identified by these studies.

With the recent development of high-density single nucleotide polymorphism (SNP) genotyping tools, it is now possible to perform large scale investigations on the genetic basis of complex traits. Genome Wide Association Studies (GWAS) with high density SNPs provides a tool to dissect the genetic architecture underlying traits by not only Identifying the QTLs and genes affecting a trait, but localising the specific SNPs with more precision. The high statistical power of detection associated with high density SNPs is due to the higher average linkage disequilibrium (LD) between SNPs. Moreover, more coverage of the genome is achieved (Khatkar & Nicholas 2008).

Objective of the study

The objective of this study was to perform a full GWAS using high density SNPs to identify with high reliability the loci that are linked to bone strength in laying hens.

MATERIAL AND METHODS

Genetic stock and housing

Birds used for the study were a pedigree population of Lohmann LSL white hens. The birds were hatched over three weeks and randomly assigned to three different houses. Hens were fed a standard commercial layer diet, with the diet and water provided *ad libitum*.

Methodology for measuring phenotype

The phenotype in the study was tibia breaking strength. The breaking strength was determined by a three-point destructive bending tests, using a JJ Lloyd LRX50 materials testing machine running the software package Nexygen 2.0 (http://www.chatillon.com) and fitted with a 2500N load cell. The bending jig consists of two 10 mm diameter steel bar supports, 30 mm apart at centre, and a 10 mm diameter cross head which approaches at 30 mm/min. Breaking strength was determined as the maximum load achieved before failure, and the failure point was set at a load which was 30% of the maximum. Stiffness was calculated from the load/displacement curve and was a measure of the bone's resistance to bending.

Data

2000 birds were initially phenotyped for tibial breaking strength using the procedure described above. The tibia bone on the right leg of the birds were sampled after slaughter. Birds which had laid less than 200 eggs in the production cycle and those that had laid less than 9 in the three weeks prior to measurement were removed, with approximately 1600 birds left. The residuals for tibial strength were then calculated by fitting body weight in a linear regression model and birds which had high leverages were also removed. Subsequently, the remaining birds were sorted based on the residuals and the top and bottom 500 were selected for the study. It was verified that there was no significant difference in body weight and egg production between the birds in the top 500 and the birds in the bottom 500.

Genotyping and quality control

The assay used for the genotyping was the Affymetrix GeneChip platform using the GeneTitan system which had 580,961 SNPs across chromosomes 1 through 28 and some unassigned. 34,841 SNPs were removed due to unknown chromosome prior to quality control. The genotype data were then subjected to a series of quality control checks using the

procedure implemented in the GenABEL R package (Aulchenko et al. 2007). SNPs with low minor allele frequency (< 1%), SNPs with low call rates (<90%), SNPs out of Hardy Weinberg equilibrium (p < 1e-12), birds with low call rates (<95%) and birds with too high autosomal heterozygosity (≥ 0.4) were removed. In total, 232,021 markers and 752 birds passed all criteria and were used for the subsequent GWAS. For the analysis, each SNP was coded by a pair of characters such as AG, and the allele A is the reference allele. For the test statistic however, the genotypes were converted to numeric. In this case, SNPs that were homozygote for the reference allele were coded as 0, the heterozygote SNPs were coded as 1 and SNPs that were homozygote for the non-reference allele were coded as 2.

Statistical analyses

An initial analysis was carried out using a linear model (analysis of covariance) to determine the possible covariates and fixed effects that could be confounded with SNP effects on tibia breaking strength. Body weight (*p*. value = 2e-16) and total egg production (*p*.value = 0.00228) were the two significant covariates identified with week of hatch (*p*.value = 8.44e-06) as a significant fixed effect. These covariates and the fixed factor were taken into account in all subsequent analyses.

Because the individuals in this study were from the same pure line population with high degree of genetic relatedness, there will be a confounding effect of the pedigree which can inflate the test statistics if a standard score test for association is used. Instead, the so-called mixed polygenic models approach was adopted in this study. Specifically, the family-based score test for association (FASTA) as put forward by Chen & Abecasis (2007) and implemented in the GenABEL R package (Aulchenko et al. 2007) was used. The FASTA approach consist of two steps: First, a mixed polygenic model is run which takes into account the genetic relationship between individuals in the study (in this case the genomic kingship matrix):

$$Y = \mu + G + e$$

where μ is the intercept (mean trait value), G is the contribution of genes to the trait value and e the residuals. This model yields the maximum likelihood estimates (MLEs) for the proportion of variance explained by genes ($\hat{\sigma}_G^2$), the error variance ($\hat{\sigma}_e^2$), fitted mean value $\hat{\mu}$ and the heritability (h^2). The model could also be modified to include possible covariates such as body weight, egg production etc. the model then becomes:

$$Y = \mu + \sum_i \beta_i . C_i + G + e$$

where C_i is a vector with the i^{th} covariate and β_i, is the coefficient of regression of the trait onto the covariate. SNP effects were also estimated at this stage by fitting each SNP at a time as a covariate in the model. However, to determine whether or not a particular SNP has a significant effect on the trait, the MLEs for the variance components (Not SNP effects) from the genetic model above were combined in a FASTA test statistics according to Chen & Abecasis (2007) as follows:

$$T_F^2 = \frac{((g - E[g])^T . (\phi . \hat{\sigma}_G^2 + I . \hat{\sigma}_e^2)^{-1} . (Y - \hat{\mu}))^2}{(g - E[g])^T . (\phi . \hat{\sigma}_G^2 + I . \hat{\sigma}_e^2)^{-1} . (g - E[g])}$$

where g is a vector containing individual genotype scores for a particular SNP (in this case the SNP being tested, $E[g]$ is a vector containing identical elements that equals 2F where F is the frequency of the A allele at the locus/SNP being tested and ϕ is the genomic kingship matrix. Y is a vector with phenotypic records and μ is the overall population mean. From the FASTA equation above, it is not immediately clear where the SNP effects are accommodated. This is because, unlike other tests such as the Wald or Likelihood ratio test, the FASTA test (being a score test) does not require actual estimates of the information under the alternative hypothesis which in this case are the solutions from the mixed polygenic model. In other words, with score test, the model estimated does not include the parameters of interest. So instead of using likelihood estimates of SNP effects from the polygenic model, the FASTA test takes into account the raw genotype and phenotype information. The score test is also very suitable for GWAS because the test is very powerful when the actual value of a parameter is close to the value under the null hypothesis, which is the case for many of the SNPs.

The FASTA procedure results in unbiased estimates of SNP effects and correct *p*-values. A genome-wide significance threshold was set as a corrected *p*-value of 10e-5. FASTA test statistic follows a chi-square distribution with one degree of freedom if the pedigree is complete and 100% correct. Because this is usually not the case, genomic control (Yang et al. 2011) was further applied to correct for possible inflations of the residuals, hence the choice of corrected *p*-values as against the standard *p*-values.

Defining QTL regions

Quantitative trait loci were defined surrounding each of the significant SNPs identified based on the local LD structure around the SNPs. For each significant SNP, a pairwise LD determined by r^2 was calculated between itself and all other SNPs within 5 mb upstream of its position and 5 mb downstream of its position using CGmisc (Kierczak et al. 2015), an R package that enables advanced analysis and visualization of GWAS data/results. With CGmisc (Kierczak et al. 2015), it was possible to graphically see the LD between an index SNP and the SNPs in its vicinity. A cut off for r^2 was set at 0.6 and any SNP whose LD with the significant SNP equals or exceeds the threshold and which was furthest upstream of the significant SNP was set as the start of the QTL and the SNP furthest downstream was set as the end of the QTL. QTLs whose positions in the genome overlapped even partially were combined into a single QTL region with the maximum and the minimum positions set as boundaries. Subsequently, these QTLs were examined to identify the genes within their boundaries, if any.

Gene set enrichment analysis

The genes identified within QTL regions were subjected to a gene set enrichment analysis using DAVID (Huang et al. 2009b; Huang et al. 2009a). The genes were also individually analyzed for possible singular functions related to bone strength.

To avoid the omission of genes which may play important roles in bone strength but are located outside the QTL regions, the genome wide significance threshold was lowered to an arbitrary figure of 0.0004. The original FASTA result was then re-checked to identify the SNPs that reached the new threshold. Using CGmisc (Kierczak et al. 2015) and the UCSC genome browser, the position of the significant SNPs were checked to see if they are located within any gene. Furthermore, a 2 mb region was defined around each of the significant SNPs, 1 mb upstream and 1 mb downstream of the significant SNP. Genes that were identified within these regions were included in the list for further gene set enrichment analysis. Special emphasis was placed on identifying the common pathways for the genes.

RESULTS

Preliminary analysis

The phenotype tibial breaking strength (in newton) had minimum and maximum values of 108.6 and 367.6 respectively and a mean value of 209.5. The standard deviation was 50.5 with a coefficient of variation of 0.24. Two covariates were identified that had significant effect ($\alpha = 0.05$) on tibia breaking strength. These were body weight (*p*. value = 2e-16) and total egg production (*p*.value = 0.00228). Week of hatch as fixed factor also showed a significant effect on tibia breaking strength (*p*.value = 8.44e-06)

Genome wide associations

After quality control, a total of 232,021 SNP markers and 752 individuals were retained and were used to estimate genome wide associations. A breakdown of the number of SNPs per chromosome left after quality control is presented in Table 1.

Table 1: Number of SNPs per chromosome

Chromosome	No. of SNPs	Chromosome	No. of SNPs	Chromosome	No. of SNPs
1	43024	11	6399	21	3979
2	26599	12	5201	22	1761
3	22625	13	3455	23	3095
4	20621	14	6089	24	3356
5	13883	15	2859	25	514
6	11025	16	171	26	2033
7	10787	17	3198	27	2336
8	8468	18	3637	28	2098
9	8888	19	3885		
10	8496	20	3539		

In the first step of the FASTA analysis, MLEs were obtained for heritability and genetic variance for tibial breaking strength. The trait showed a very high heritability of 0.55 and a genetic standard deviation of 46.6. In the second step, SNP effects with their standard errors, 1 degree of freedom chi-square test for association, standard p-values and corrected p-values (after genomic control) were estimated and are presented in the supplementary table. From the result, a total of 52 SNPs reached the genome wide threshold of 10e-5. These SNPs were spread across chromosome 1 (2 SNPs), chromosome 3 (29 SNPs), chromosome 8 (20 SNPs) and chromosome 16 (1 SNP) (**figure 1**). The top SNP identified in the study was subsequently fitted as a covariate in a polygenic model and the heritability of the trait was re-estimated. The value obtained for heritability after this procedure was 0.53, which means that about 2% of the variation in tibial breaking strength phenotype is explained by allelic variation at this locus alone.

An LD analysis was carried out in the regions containing the significant SNPs, which was done to establish if the significant SNPs were in LD each other and with other SNPs in the regions. On chromosome 1, the two significant SNPs were not in LD with each other, suggesting that there are two separate regions of interest on the chromosome. It is expected that they will not be in LD given that they are spaced far apart (20 mb distance). On a closer examination, it was found that not only were these SNPs far apart and not in LD with each other, they were also not in LD with any other SNP in their vicinity (**figure 2**). Because they are singular SNPs and not in LD with other SNPs, their minor allele frequencies (MAF) were checked and it was found that they both had very low MAFs which barely passed the cut-off criteria for MAF of 0.01 as implemented during quality control. Their MAFs were 0.02582 and 0.01962 respectively.

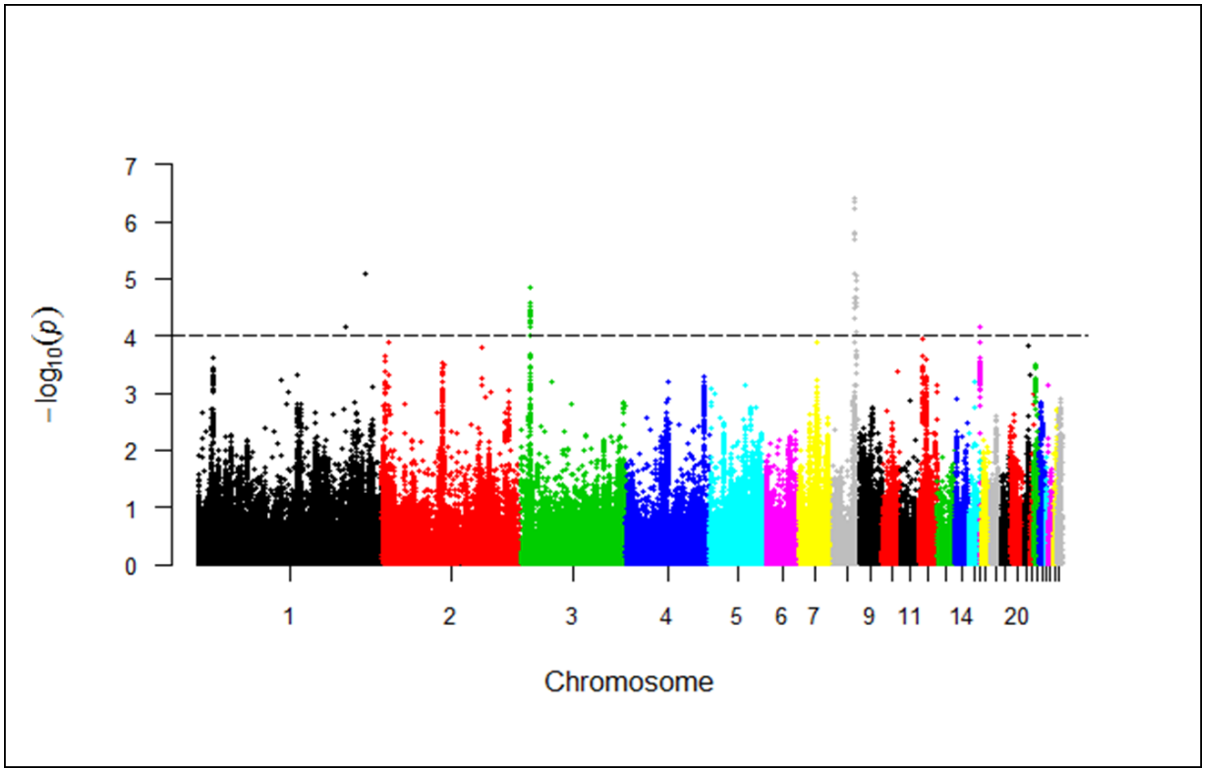

Figure 1: Manhattan plot of genome wide associations for tibia breaking strength in laying hens*. The -log$_{10}$ of corrected p-values is shown for each SNP (y-axis). The genome wide threshold is indicated by a horizontal dashed line*

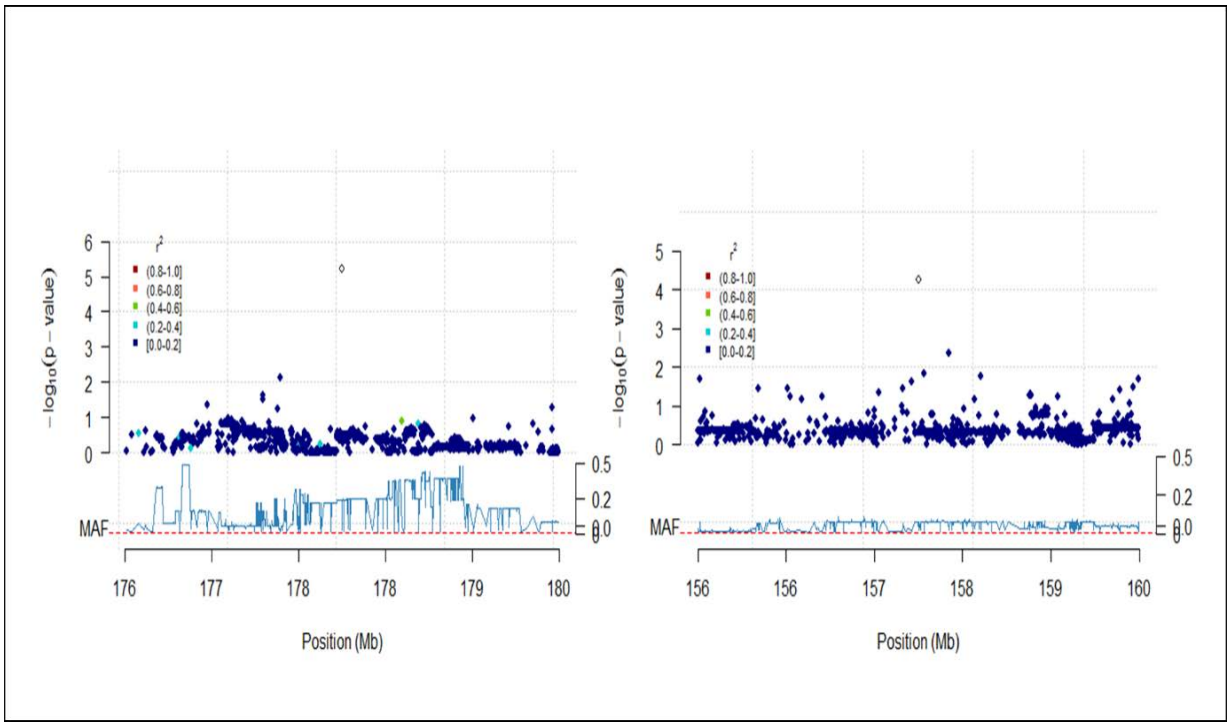

Figure 2: LD plots showing the local LD structure around the two significant SNPs on chromosome 1

Chromosome 3 had the highest number of significant SNPs with a total of 29 that passed the genome wide significance threshold. The result showed that these SNPs were very close to each other and were all within a 1 mb region. LD analysis also showed they were in high LD with each other and with other non-significant SNPs in the region (**figure 3**).

The strongest association signal was observed within the significant region on chromosome 8. This region had a total of 20 SNPs that reached the genome wide significance threshold, all of which were close to each other and were within a 1 mb region (**figure 4**). The most significant SNP in the region had a corrected p.value of 4e-7. Chromosome 16 had the fewest number of markers compared to the other chromosomes (**Table 1**). This chromosome had only a single SNP that reached the genome wide significance threshold, but the SNP was in high LD with other SNPs in its vicinity (**figure 5**).

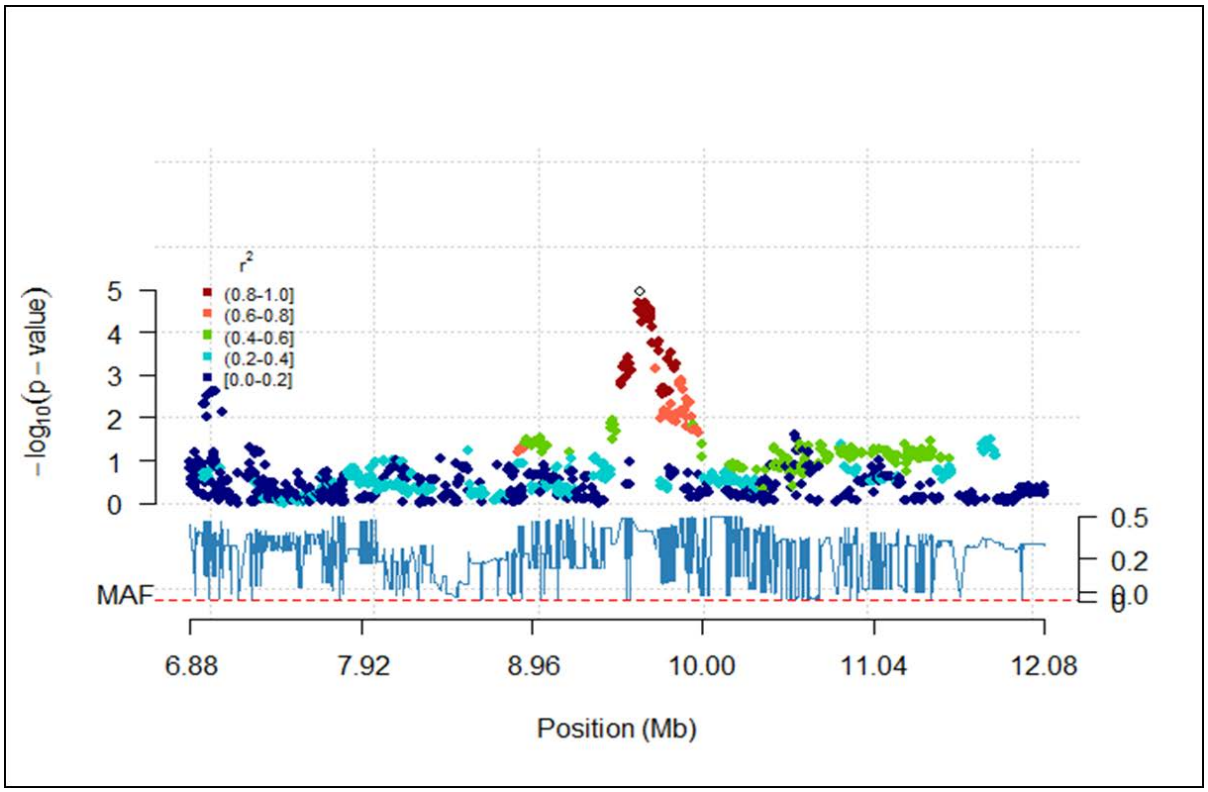

Figure 3: LD plot showing the LD structure around the most significant SNP on chromosome 3

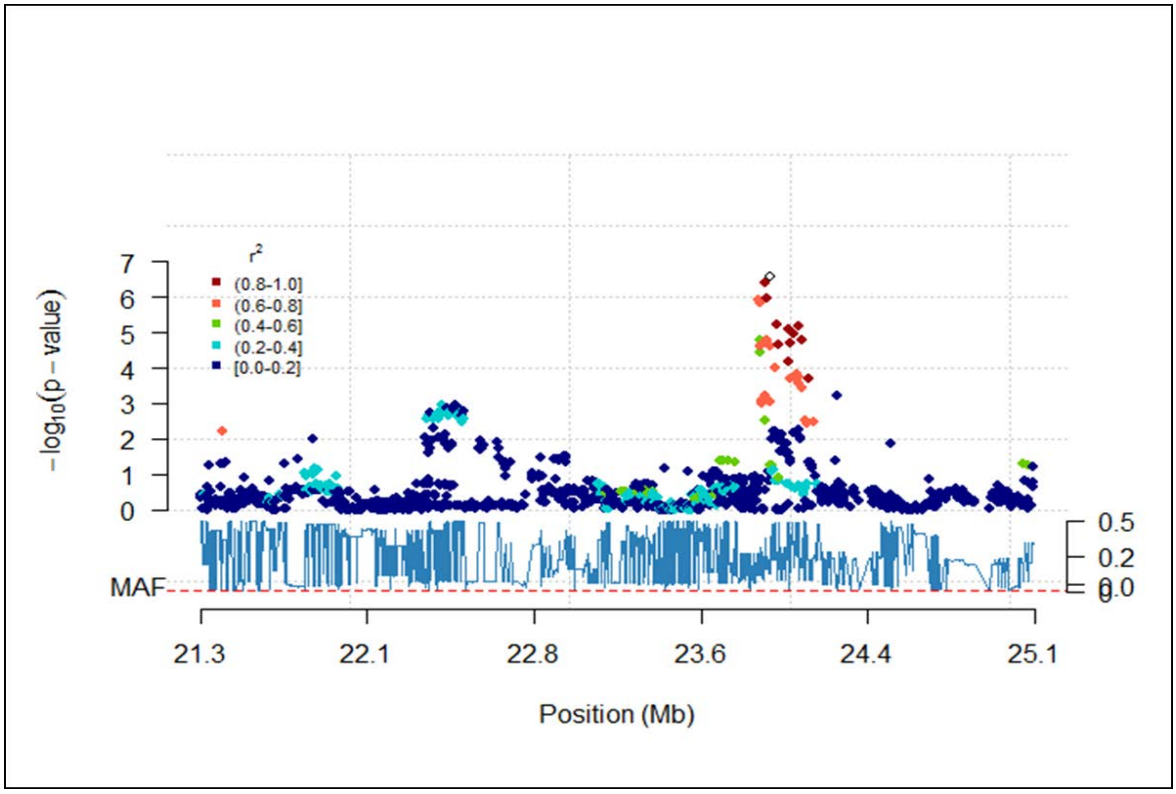

Figure 4: LD plot showing the local LD structure around the most significant SNP on chromosome 8

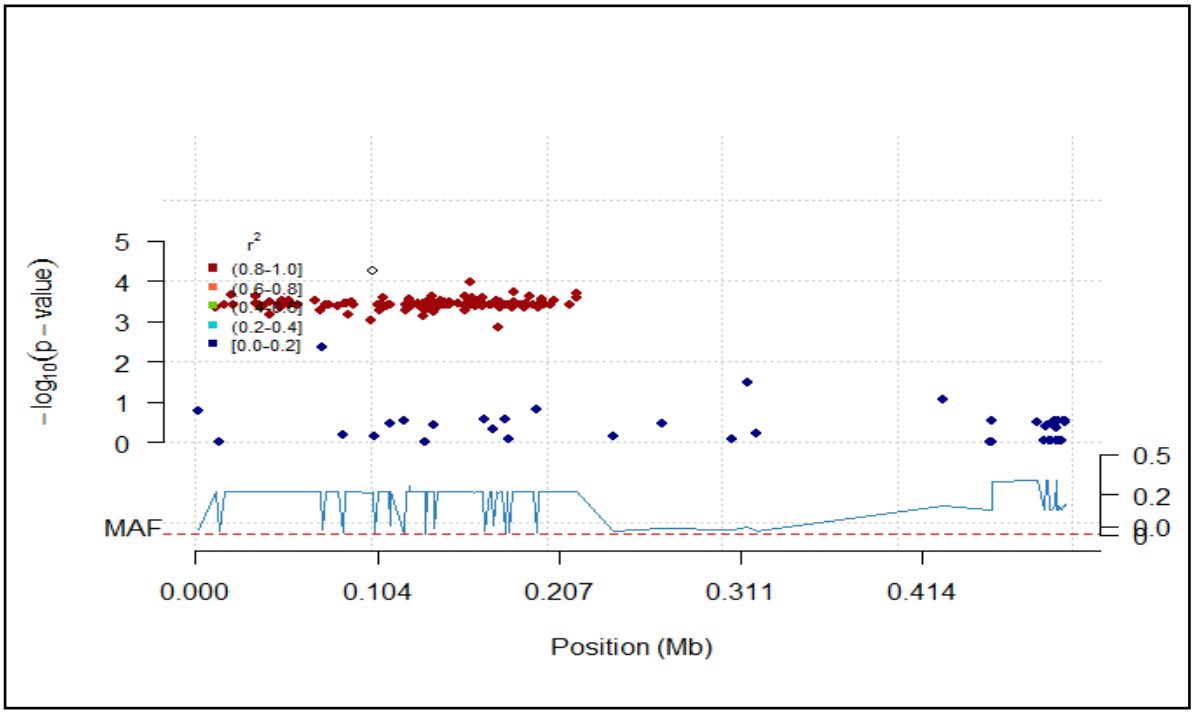

Figure 5: LD plot showing the local LD structure around the significant SNP on chromosome 16

The effect of the most significant SNP (AX-77092894, chr.8) on the phenotype was investigated to see if different allele combinations for this SNP results in observed differences in the phenotype (**figure 6**). The result clearly indicate that individuals with A/A at this locus have higher average bone strength than individuals with A/G or G/G at the same locus. The difference in average bone strength between A/A individuals and G/G individuals for example is approximately 40 newton.

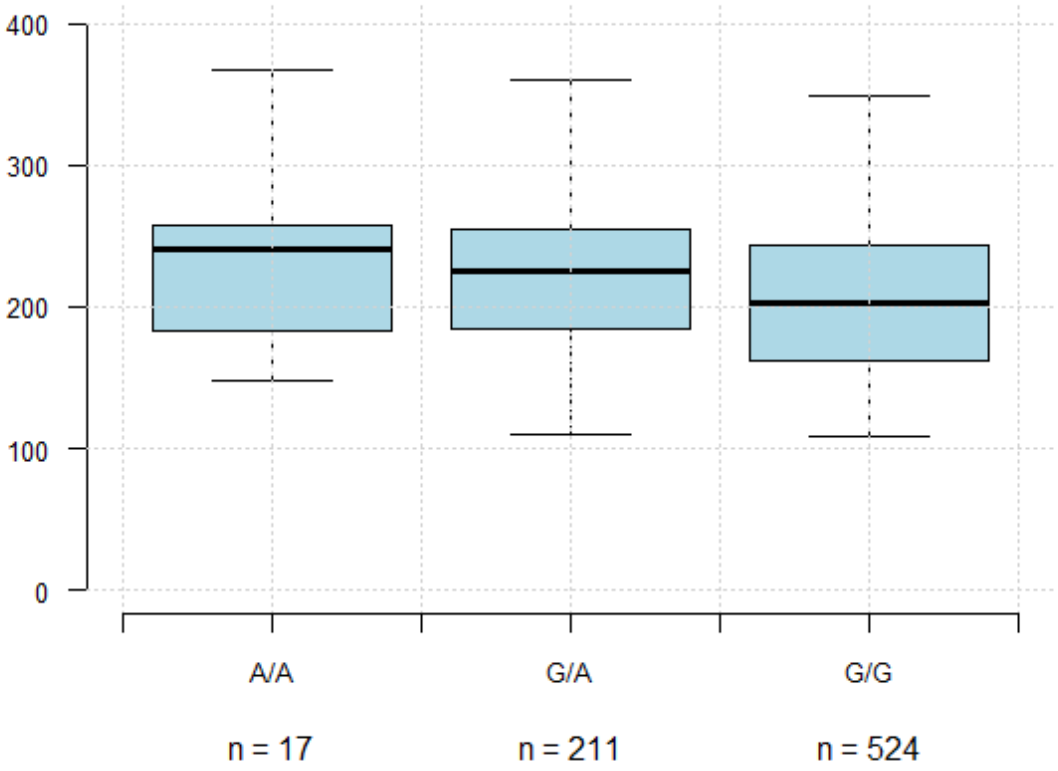

Figure 6: Boxplot showing the effect of the most significant SNP (AX-77092894, chr.8) on tibial breaking strength phenotype

QTL regions defined

After defining the local LD range for each significant SNP on chromosome 3, all the LD ranges nicely overlapped into a single QTL region. The same was observed for the significant SNPs on chromosome 8 which also formed a single QTL region. The LD range around the significant SNP on chromosome 16 also represented a QTL region. Taking into account the two separate SNPs on chromosome 1, there were in total, 5 distinct QTLs for bone strength that were identified in the study (**Table 2**)

Table 2: QTL found in the chicken genome associated with tibia breaking strength.

chromosome	QTL start (bp)	QTL end (bp)	QTL length (bp)	No. of significant SNPs in QTL
1	178054319	178054319	1	1
1	157504202	157504202	1	1
3	8878928	9976543	1097616	29
8	21420506	24110421	2689916	20
16	11466	217707	206242	1

Genes identified within the QTL regions

The positions of the two significant SNPs on chromosome 1 were checked in the UCSC genome browser (Kent et al. 2002) to see if they are located within any gene. The result showed that they are both not located within any gene. The QTL region on chromosome 3 was also examined for the presence of genes. The search turned up a number of known genes and some Ensembl predicted genes (**Table 3**). Some of the genes located within this region code for proteins that are yet to be characterised while some of the genes code for proteins that are involved in processes unrelated to bone strength. Some of the genes however have functions that are related to skeletal development.

Table 3: QTL regions and the genes within them

Chr	QTL start (bp)	QTL end (bp)	QTL length	No. of SNPs	No. of Genes	Gene names*
1	178054319	178054319	1	1	0	
1	157504202	157504202	1	1	0	
3	8878928	9976543	1097616	29	23	COMMD1 B3GNT2 ENSGALT00000039588.2 ENSGALT00000014512.3 TMEM17 EHBP1 ENSGALT00000042631.1 WDPCP MDH1 UGP2 VPS54 PELI1 ENSGALT00000043571.1 ENSGALT00000042828.1 LGALSL AFTPH SERTAD2 SLC1A4 CEP68 RAB1A ACTR2 SPRED2 gga-mir-6711
8	21420506	24110421	2689916	20	51	SLC5A9 SPATA6 gga-mir-1809 BEND5 ELAVL4 ENSGALT00000042744.1 FAF1 CDKN2C RNF11 TTC39A EPS15 OSBPL9 gga-mir-1562 NRD1 RAB3B TXNDC12 BTF3L4 ZFYVE9 CC2D1B ORC1 PRPF38A

						ZCCHC11 GPX7 FAM159A ENSGALT00000025940.3 gga-mir-6549 SELRC1 ZYG11B ECHDC2 SCP2 PODN SLC1A7 gga-mir-6623 CPT2 C1orf123 gga-mir-1675 MAGOH2 LRP8 DMRTB1 GLIS1 NDC1 YIPF1 DIO1 LRRC42 LDLRAD1 TMEM59 TCEANC2 CDCP2 CYB5RL MRPL37 SSBP3
16	11466	217707	206242	1	28	ENSGALT00000000238.4 CYP21A2 CenpA C4 BFIV21 TAP2 TAP1 BF1 DMB2 ENSGALT00000041213.2 DMA BRD2 TPN BLB1 Blec2 BG1 B-BTN2 GNB2L1 SNORD95 TRIM27.2

	TRIM27.1 HEP21 TRIM7.1 IL4I1 ENSGALT00000043587.1 ENSGALT00000045620.1 TRIM7.2 ZNF692

The Ensembl gene identifier is provided where a gene name was not available.

A number of known genes and Ensembl predicted genes were also annotated within the QTL region on chromosome 8 and also within the QTL region on chromosome 16 (**Table 2**).

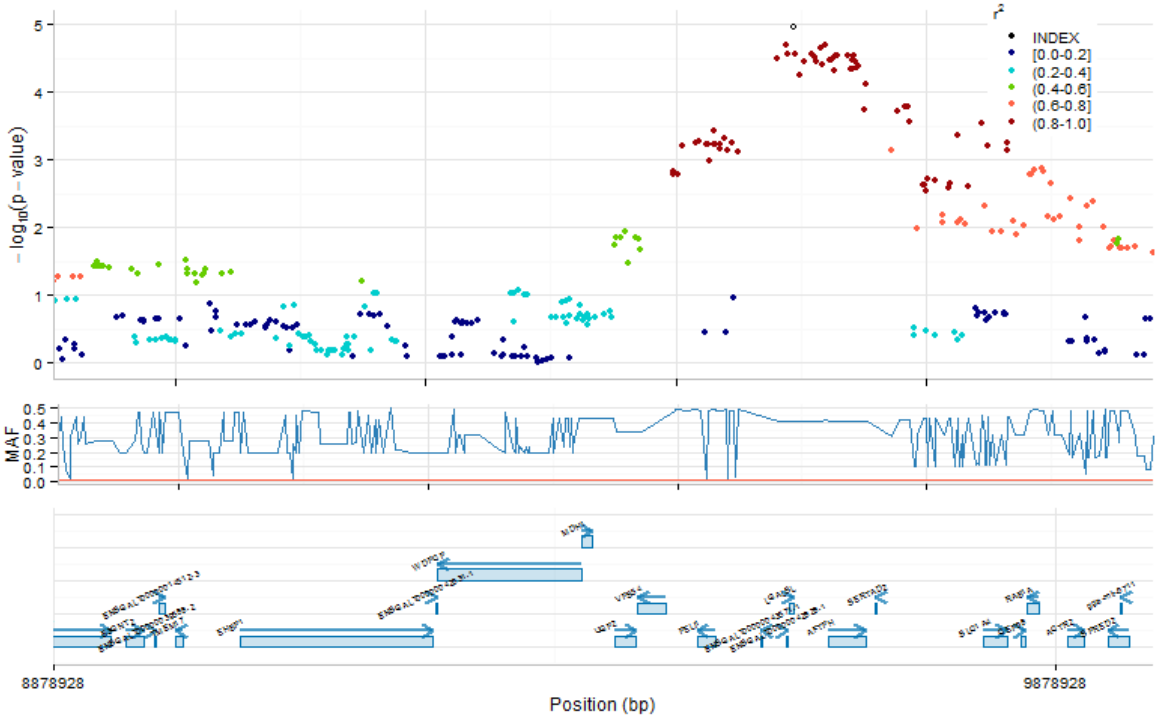

Figure 7: A plot of the QTL region on chromosome 3 with the genes and their positions plotted in the lower panel

Gene set enrichment analysis after lowering the genome wide significance threshold

After the genome wide significance threshold was lowered to 0.0004, a total of 121 SNPs across 9 different chromosomes were identified as significant. A 2 mb region was defined around each of these significant SNPs and genes within these regions were identified. The genes are presented in table 4.

Table 4: Genes within the 2 mb regions surrounding each significant SNP

Chr1	Chr2	Chr3	Chr7	Chr8	Chr12	Chr16	Chr21	Chr23
Fgf9	DPP6	MSH6	LRP2	RNF11	DAG1	CD1A1	NOC2L	
MICU2	PAXIP1	FBXO11	DHRS9	TTC39A	MST1	CD1.2	SAMD11	
ZDHHC20	HTR5A	FOXN2	ABCB11	EPS15	APEH	LTB4R2	RERE	
SKA3	INSIG1	PPP1R21	G6PC2	OSBPL9	BSN	CYP21A2	ENO1	
chSAP18	EN2	COMMD1	SPC25	gga-mir-1562	TRAIP	CenpA	CA6	
LATS2	RBM33	B3GNT2	CERS6	NRD1	gga-mir-1678	C4	GPR157	
XPO4	SHH	TMEM17	gga-mir-1733	RAB3B	CAMKV	BFIV21	gga-mir-34a	
N6AMT2	RNF32	EHBP1	STK39	TXNDC12	MST1R	TAP2	H6PD	
IFT88	LMBR1	WDPCP	B3GALT1	BTF3L4	MON1A	TAP1	SPSB1	
CRYL1	NOM1	MDH1	gga-mir-1591	ZFYVE9	RBM6	BF1	SLC25A33	
GJB6	MNX1	UGP2	SCN9A	CC2D1B	RAD54L2	DMB2	TMEM201	FOXO6
GJA3	UBE3C	VPS54	TTC21B	ORC1	FANCD2	DMA	PIK3CD	SCMH1
ZMYM2	DNAJB6	PELI1	GALNT3	PRPF38A	EMC3	BRD2	calsyntenin-1	
PSPC1	PTPRN2	LGALSL	CSRNP3	ZCCHC11	USP4	TPN	CTNNBIP1	
gga-mir-6641	gga-mir-153	AFTPH	SCN2A	GPX7	GPX1	BLB1	LZIC	
MPHOSPH8	DSG2	SERTAD2	SCN3A	FAM159A	cRhoA	Blec1	NMNAT1	
CENPJ	TTR	SLC1A4	SLC38A11	gga-mir-6549	TEX264	Blec2	RBP7	
RNF17	B4GALT6	CEP68	COBLL1	SELRC1	GRM2	BG1	UBE4B	
[2Fe-2S]	TRAPPC8	RAB1A	GRB14	ZYG11B	ALAS1	B-BTN2	KIF1B	
RDX	RNF138	ACTR2	FIGN	ECHDC2	TWF2	GNB2L1	uc_338	
ZC3H12C	MEP1B	SPRED2	KCNH7	SCP2	PPM1M	SNORD95	PGD	
KLHL1	GAREM	gga-mir-6711		PODN	WDR82	TRIM27.2	APITD1	
PCDH9	KLHL14	MEIS1		SLC1A7	gga-let-7g	TRIM27.1	DFFA	
PRKAR2B	gga-mir-6707			gga-mir-6623	GLYCTK	HEP21	PEX14	
HBP1	CCDC178			CPT2	gga-mir-135a-1	TRIM7.1	CASZ1	
COG5	ASXL3			C1orf123	DNAH1	IL4I1	TARDBP	
GPR22	DTNA			gga-mir-1675	HAUS3	TRIM7.2	MASP2	
DUS4L	VIPR1			MAGOH2	BAP1	ZNF692	SS2	
gga-mir-2126	gga-mir-1662			LRP8	SEMA3G	CLEC2L	SRM	
BCAP29	SEC22C			DMRTB1	ABHD14A	KIFC1	EXOSC10	
SLC26A4	NKTR			GLIS1	GPR61	B-G	MTOR	
CBLL1	ZBTB47			NDC1	PARP3		ANGPTL7	
SLC26A3	KLHL40			YIPF1	RBM5		UBIAD1	
DLD	HHATL			DIO1	Sema3F		PLEKHM2	
LAMB1	CCDC13			LRRC42	GNAT1		TMEM82	
LAMB4	HIGD1A			LDLRAD1	GNAI2		FBLIM1	
ERGIC2	ACKR2			TMEM59	SLC38A3		SPEN	
LRRK2	OBSCN			TCEANC2	SEMA3B		ZBTB17	
SLC2A13	IBA57			CDCP2	OAS*A		HSPB7	
C12orf40	GJC2			CYB5RL	HYAL3		EphA2	
ABCD2	GUK1			MRPL37	SHISA5		FBXO42	
KIF21A	MRPL55			SSBP3	ATRIP		C21H1ORF144	
CPNE8	C1orf35			ACOT11	IPPK		NECAP2	
ALG10	ARF1			TTC4	CENPP		MFAP2	
TPRKB	Wnt3a			PARS2	ECM2		ATP13A2	
GTSE1	gga-mir-6713			C1orf177	ASPN		Pax-7	
TRMU	WNT9a			DHCR24	OMD		ALDH4A1	
CELSR1	SNAP47			TMEM61	OGN		UBR4	
GRAMD4	JMJD4			BSND	NOL8		gga-mir-6626	
CERK	ALS2CL			PCSK9	IARS		EMC1	
TBC1D22A	TMIE			USP24	SNORA84		MRTO4	
	MYL3			PPAP2B	ATP2B2		PQLC2	
	PTH1R			PRKAA2	SLC6A11		CAPZB	
					SLC6A1		MINOS1	
					HRH1		NBL1	
					ATG7		HTR6	
							TMCO4	

gga-mir-460a	C1orf168	CDHR4	RNF186
NEDD9	C8A	UBA7	OTUD3
ELOVL2	C8B	AMT	PLA2G2E
SYCP2L	Dab1	CACNA1D	PLA2G5
GCM2		CHDH	PLA2
MAK		ACTR8	UBXN10
GCNT2		SELK	DDX19B
TFAP2A		CACNA2D3	MUL1
OFCC1		LRTM1	FAM43B
SLC35B3		WNT5A	CDA
BLOC1S5		ERC2	PINK1
TXNDC5		CCDC66	
gga-mir-6571		FAM208A	
BMP6		ARHGEF3	
DSP		IL17RD	
RIOK1		HESX1	
SSR1		APPL1	
RREB1		ASB14	
LY86		PDE12	
F13A1		ARF4	
FARS2		DENND6A	
LYRM4		SLMAP	
RPP40		gga-mir-1787	
ECI2		gga-mir-1783	
PRPF4B		FLNB	
PXDC1		DNASE1L3	
SLC22A23		ABHD6	
TUBB2B		IP6K2	
BPHL		NCKIPSD	
RIPK1		CELSR3	
NQO2		SLC26A6	
DUSP22		UQCRC1	
IRF-4		COL7A1	
EXOC2		COPG1	
SNORD77		C12H3ORF37	
GMDS		RAB7A	
		RNF123	
		AMIGO3	
		GMPPB	
		IP6K1	

Despite the high number of genes identified, only a few of the identified genes were seen to be involved together in the different pathways. DAVID reported four significant KEGG pathways with each pathway involving 3 to 5 of the genes in table 4. The pathways are:

Table 5: significant KEGG pathways involving the identified genes

pathway	Listed genes	% of genes	P-Value	Benjamini
Glycine, serine and threonine metabolism	5	1.6	3.70E-03	2.60E-01
Hedgehog signaling pathway	5	1.6	2.30E-02	6.20E-01
Glutathione metabolism	4	1.2	5.00E-02	7.50E-01
Glycosphingolipid biosynthesis	3	0.9	7.90E-02	8.20E-01

DISCUSSION

In this study, we have shown that bone strength is indeed highly influenced by genetics in addition to environmental factors. The phenotype tibial breaking strength as a representation of bone strength was highly variable in the population we studied. Although the observations were censored, in that only the top and bottom individuals in terms of bone strength were selected, the difference between the observed minimum and maximum value is a clear indication of the amount of variation that exist for this trait. In the study of Bishop et al. (2000), they reported a heritability for tibial strength to be 0.45 which is lower than the heritability we found for tibial breaking strength in our study (0.55). The reason for this higher heritability may be because we used high density markers and genomic kinship matrix in our estimation which is able to capture more genetic variation than when using a classical BLUP and pedigree based kingship matrix (Meuwissen 2007). It may also be that the heritability is higher because the population on which we performed our estimation is a pre-selected population. Individuals were included in the study based on their phenotypic value and therefore the heritability of the trait in this case may not be a true representation of the heritability in an unselected population. In whatever case, it is clear that the heritability is higher than for most studied traits, which means that the trait can easily be changed through genetic selection in a relatively shorter period of time.

This study unlike previous studies, utilised a substantially larger number of SNPs which resulted in higher resolution and increased power/accuracy of detecting QTLs linked to bone strength. This is the first time to the best of our knowledge that high density SNP markers are utilised for a GWAS study on bone strength in laying hens. The genome wide significance threshold of 10e-5 was arbitrarily assigned and not based on Bonferroni correction. Several studies have shown the Bonferroni correction for multiple comparisons to be overly conservative (Gao et al. 2010; Johnson et al. 2010) and more likely to inflate the probability of type 2 error (incorrectly assigning non-significance to a statistical test). This is because of the assumption of the Bonferroni procedure that the SNPs in the study are independent and not in LD with each other which doesn't hold in reality. An alternative to Bonferroni correction would have been to implement the permutation test which is considered to be the Gold standard in the scientific community. But apart from the fact that it is computationally demanding especially when there are thousands of SNPs in the study, as a disadvantage of the FASTA procedure, permutation test cannot be applied to obtain genome wide significance because the data structure is not exchangeable.

We therefore based our arbitrary genome wide significance by looking at the commonly used p.value for GWAS applied to livestock which is usually between 10e-4 and 10e-6. Even in human studies, the genome wide threshold is not very much lower than this. In a GWAS applied to a European decent population for example, the standard p.value is around 5e-8, but this is proposed to correct for about 1-2 million independent tests (McCarthy & Abecasis 2008; Risch & Merikangas 1996). It must be emphasised that in this study, the correction is for much lower number of tests (232,021) compared to the number of tests in human studies.

The study identified a total of 52 SNPs that reached or exceeded the genome wide threshold of 10e-5. These SNPs were spread across chromosome 8 (20 SNPs), chromosome 3 (29 SNPs), chromosome 1 (2 SNPs) and chromosome 16 (1 SNP). The study also found 5 distinct QTLs linked to bone strength. Because the two significant SNPs on chromosome 1 were not in LD with each other given their distance apart (20mb), they were considered to be separate QTLs. The SNPs were also not in LD with other SNPs in their surrounding and were not located within any known gene. Dunn & Fleming (2007) in their study found a significant QTL for osteoporosis on chromosome one. The position of the QTL they found was 370cM on chromosome 1, 1 megabase upstream of the RUNX1 gene. This position in base pairs corresponds to 108,473,589 which is 65 megabase upstream of the first QTL we found on chromosome 1 and 49 megabase upstream of the second QTL on chromosome 1 (Table 2). It should be noted however that their annotation was based on the galGal3 chicken assembly, while our annotation was based on the galGal4 assembly.

There was a relatively large QTL detected on chromosome 3 ranging from 8878928bs to 9976543bs. This QTL had a number of genes annotated within its boundaries (see table 3). Melissa et al. (2005) found several suggestive QTLs linked to bone traits in laying hens with some of the suggestive QTLs found on chromosome 3. These suggestive QTLs were however not significant after adjusting for the variation in body weight and egg production. Genes identified within this QTL perform several functions but the ones are related to bone strength are:

Transmembrane Protein 17 (TMEM17): This gene is required for ciliogenesis and sonic hedgehog/SHH signalling, with both processes playing critical roles in skeletal development in vertebrates (Goetz & Anderson 2010; Nosavanh et al. 2015).

Actin-Related Protein 2 (**ACTR2**): A very important biological process involving this gene is cilium assembly or ciliogenesis. Cilia as pointed out above play important roles in skeletal development (Goetz & Anderson 2010)

Solute Carrier Family 1 (Glutamate/Neutral Amino Acid Transporter), Member 4 (**SLC1A4**): This gene has been shown to have some implications in the proper functioning of skeletal muscles (Kanai & Hediger 2003).

(4) **WD Repeat Containing Planar Cell Polarity Effector (WDPCP)**: This gene also plays a role in ciliogenesis (Viguet-Carrin et al. 2006).

The strongest association was detected within the QTL region on chromosome 8. It was surprising however that most of the genes identified within this region were participating in other functions unrelated to bone strength, mostly immunity functions. Genes whose function are related to bone strength are:

Podocan (PODN): The human ortholog of this gene has been shown to be involved in collagen binding and development. Collagen on the other hand plays an important role in bone strength (Viguet-Carrin et al. 2006).

Single Stranded DNA Binding Protein 3 (SSBP3): This gene may be involved in transcription regulation of the alpha 2(I) collagen gene, thereby playing an indirect role in bone strength. To the best of our knowledge, this is the first study to identify a QTL on chromosome 8 related to bone traits in laying hens.

Another QTL was found on chromosome 16. There was only one gene within this region however whose function is related to bone strength. This was the **Osteoclast inhibitory lectin (BRD2)** which is required for normal bone physiology (Kartsogiannis et al. 2008). In a human study, this gene was associated with a reduction of bone mineral density in women (Pineda et al. 2008). This is also a novel QTL, given that no other study has reported a QTL on chromosome 16 linked to bone traits in laying hens.

In our attempt to identify the pathways in which potential candidate genes are involved, we lowered the genome wide significance threshold to 0.0004. Given this new threshold, several other genes were identified (Table 4) and the pathways in which these genes are involve (Table 5).

Of the pathways identified in table 5 however, only the Hedgehog signalling pathway has been linked to bone strength. This pathway together with the Wnt signalling pathway is thought to initiate osteogenic development in vertebrates through its regulation of endochondral ossification (Kronenberg 2003; Karsenty 2003), one of the two ways through which bones are produced. Genes identified in this study that are involved in the pathway are: BMP6, LRP2, SHH, WNT5A and WNT9a.

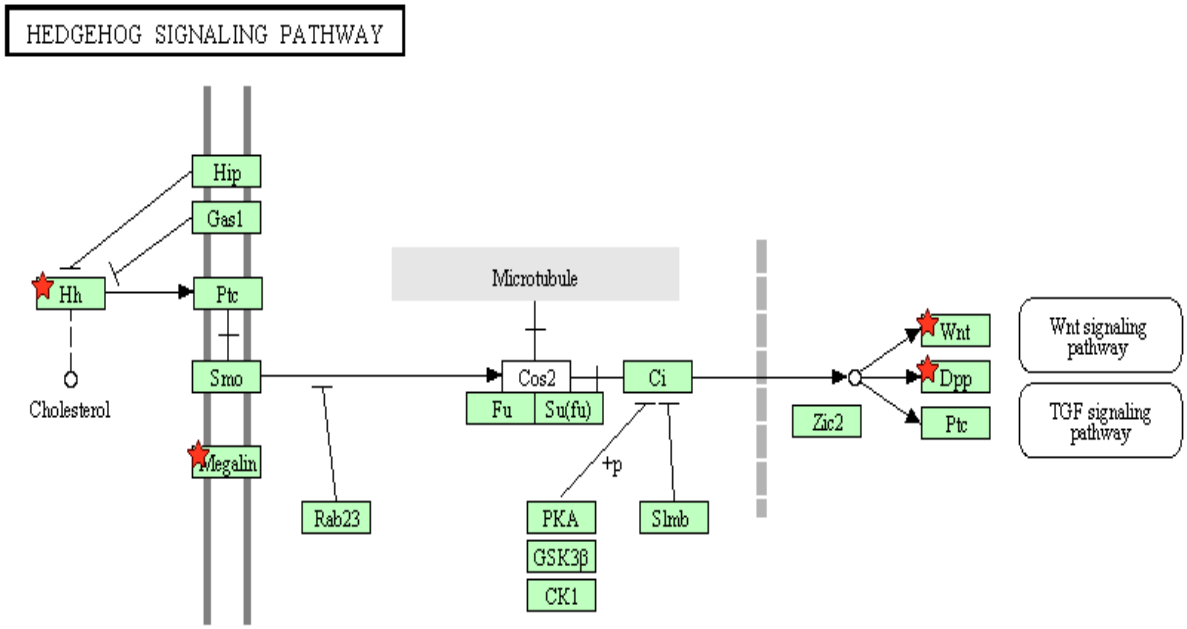

Figure 8: The Hedgehog signalling pathway (red star indicate where the genes are involved)

Other important pathways (panther pathways) were identified in the study which include the following.

Wnt signalling pathway: It is obvious from the pathway chart for Hedgehog signalling above that these pathways are involved in similar processes. In animal studies, this pathway was shown to regulate cartilage development, growth, and maintenance (Usami et al. 2015). It is also required to ensure the progression of endochondral ossification and development of axial and appendicular skeletons. The studies of Gong et al. (2001) and Kato et al. (2002) pointed out the role that the Wnt pathway plays in the regulation of bone mass. CTNNBIP1, WNT5A and WNT9A are the genes in table 4 that are involved in this pathway.

Cadherin signalling pathway: This pathway plays a role in the regulation of osteogenic differentiation and mechanotransduction (Marie et al. 2014). WNT9A, WNT5A and Celsr3 are the genes involved in this pathway.

FAS signalling pathway: In a study, increase in osteoclast numbers and activity, along with reduced bone mass was observed when FASL was conditionally knocked out in osteoblasts (Wang et al. 2015), which suggests that osteoblast-produced FASL is important in maintaining bone mass.

CONCLUSION

This study have identified loci linked to tibial breaking strength in laying hens. 52 significant SNPs, 5 distinct and novel QTLs were found across chromosome 1, 3, 8 and 16. These QTL regions had a number of promising candidate genes, some of which have been shown to participate in processes influencing bone strength in laying hens. Gene enrichment analysis revealed important pathways such as Hedgehog signalling and Wnt which are linked to bone strength and in which some of the identified genes play critical roles. The identified QTLs and the genes they encompass provides important information for genetic selection to improve bone strength and ultimately the welfare of layers. They also form the basis for future research into the genetic architecture of bone strength in laying hens.

DEDICATION

To my Lord and saviour Jesus Christ for grace, wisdom and strength. To my dad who passed on before I could finish my master program, I hope this makes you proud.

REFERENCES

Aulchenko, Y.S. et al., 2007. GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, 23(10), pp.1294–1296.

Bishop, S.C. et al., 2000. Inheritance of bone characteristics affecting osteoporosis in laying hens. *British poultry science*, 41(1), pp.33–40.

Chen, W.-M. & Abecasis, G.R., 2007. Family-based association tests for genomewide association scans. *American journal of human genetics*, 81(5), pp.913–26.

Dunn, I. & Fleming, R., 2007. A QTL for osteoporosis detected in an F2 population derived from White Leghorn chicken lines divergently selected for bone index. *Animal* ….

Fleming, R. & McCormack, H., 2006. Relationships between genetic, environmental and nutritional factors influencing osteoporosis in laying hens. *British poultry* ….

Gao, X. et al., 2010. Avoiding the high Bonferroni penalty in genome-wide association studies. *Genetic epidemiology*, 34(1), pp.100–5.

Goetz, S.C. & Anderson, K. V, 2010. The primary cilium: a signalling centre during vertebrate development. *Nature reviews. Genetics*, 11(5), pp.331–44.

Gong, Y. et al., 2001. LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development. *Cell*, 107(4), pp.513–23.

Gregory, N. & Wilkins, L., 1990. Broken bones in domestic fowls: Effect of husbandry system and stunning method in end-of-lay hens. *British Poultry* ….

Gregory, N., Wilkins, L. & Knowles, T., 1994. Incidence of bone fractures in European layers.

Gregory, N.G. & Wilkins, L.J., 1989. Broken bones in domestic fowl: handling and processing damage in end-of-lay battery hens. *British poultry science*, 30(3), pp.555–62.

Huang, D.W., Sherman, B.T. & Lempicki, R.A., 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*, 37(1), pp.1–13.

Huang, D.W., Sherman, B.T. & Lempicki, R.A., 2009b. Systematic and integrative analysis

of large gene lists using DAVID bioinformatics resources. *Nature protocols*, 4(1), pp.44–57.

Johnson, R.C. et al., 2010. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC genomics*, 11(1), p.724.

Kanai, Y. & Hediger, M.A., 2003. The glutamate and neutral amino acid transporter family: physiological and pharmacological implications. *European Journal of Pharmacology*, 479(1-3), pp.237–247.

Karsenty, G., 2003. The complexities of skeletal biology. *Nature*, 423(6937), pp.316–8.

Kartsogiannis, V. et al., 2008. Osteoclast inhibitory lectin, an immune cell product that is required for normal bone physiology in vivo. *The Journal of biological chemistry*, 283(45), pp.30850–60.

Kato, M. et al., 2002. Cbfa1-independent decrease in osteoblast proliferation, osteopenia, and persistent embryonic eye vascularization in mice deficient in Lrp5, a Wnt coreceptor. *The Journal of cell biology*, 157(2), pp.303–14.

Kent, W.J. et al., 2002. The human genome browser at UCSC. *Genome research*, 12(6), pp.996–1006.

Khatkar, M. & Nicholas, F., 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC* ….

Kierczak, M. et al., 2015. cgmisc: enhanced genome-wide association analyses and visualization. *Bioinformatics (Oxford, England)*, p.btv426–.

Kronenberg, H.M., 2003. Developmental regulation of the growth plate. *Nature*, 423(6937), pp.332–6.

MANDOUR, M.A. et al., 1989. Selection for Increased Humerus Strength of Cage-Reared Broilers. *Poultry Science*, 68(9), pp.1168–1173.

Marie, P.J., Haÿ, E. & Saidak, Z., 2014. Integrin and cadherin signaling in bone: role and potential therapeutic targets. *Trends in endocrinology and metabolism: TEM*, 25(11), pp.567–75.

McCarthy, M. & Abecasis, G., 2008. Genome-wide association studies for complex traits:

consensus, uncertainty and challenges. *Nature Reviews* ….

Melissa, A., Patricia, Y. & Diane, E., 2005. Identification of quantitative trait loci associated with bone traits and body weight in an F2 resource population of chickens. *Genet. Sel. Evol*.

Meuwissen, T., 2007. Genomic selection: marker assisted selection on a genome wide scale. *Journal of Animal Breeding and Genetics*.

Nosavanh, L. et al., 2015. Cell-autonomous activation of Hedgehog signaling inhibits brown adipose tissue development. *Proceedings of the National Academy of Sciences of the United States of America*, 112(16), pp.5069–74.

Pineda, B. et al., 2008. The Asn19Lys substitution in the osteoclast inhibitory lectin (OCIL) gene is associated with a reduction of bone mineral density in postmenopausal women. *Calcified tissue international*, 82(5), pp.348–53.

Riddell, C., 1992. Non-infectious skeletal disorders of poultry: an overview. *Poultry Science Symposium*.

Risch, N. & Merikangas, K., 1996. The future of genetic studies of complex human diseases. *Science*.

Rubin, C.-J. et al., 2007. Quantitative trait loci for BMD and bone strength in an intercross between domestic and wildtype chickens. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*, 22(3), pp.375–84.

Schreiweis, M. & Orban, J., 2003. The use of densitometry to detect differences in bone mineral density and content of live White Leghorns fed varying levels of dietary calcium. *Poultry science*.

Schreiweis, M., Orban, J. & Ledur, M., 2005. Validation of dual-energy X-ray absorptiometry in live White Leghorns. *Poultry* ….

Spencer, C. et al., 2009. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet*.

Usami, Y. et al., 2015. Wnt signaling in cartilage development and diseases: lessons from

animal studies. *Laboratory investigation; a journal of technical methods and pathology*.

Viguet-Carrin, S., Garnero, P. & Delmas, P.D., 2006. The role of collagen in bone strength. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA*, 17(3), pp.319–36.

Wang, L. et al., 2015. Osteoblast-induced osteoclast apoptosis by fas ligand/FAS pathway is required for maintenance of bone mass. *Cell death and differentiation*, 22(10), pp.1654–64.

Weber, R., 2003. Investigations of laying hen health in enriched cages as compared to conventional cages and a floor pen system. *Wiener* ….

Webster, A., 2004. Welfare implications of avian osteoporosis. *Poultry Science*.

Yang, J. et al., 2011. Genomic inflation factors under polygenic inheritance. *European journal of human genetics : EJHG*, 19(7), pp.807–12.