Genome-Wide Association Studies for Immunoglobulins in Colostrum and Serum in Chinese Holstein

Shan Lin  
China Agricultural University  
Cuncun Ke  
China Agricultural University  
Lin Liu  
Beijing Dairy Cattle Center  
Yahui Gao  
China Agricultural University  
Lingna Xu  
China Agricultural University  
Bo Han  
China Agricultural University  
Yaofeng Zhao  
China Agricultural University  
Shengli Zhang  
China Agricultural University  
Dongxiao Sun  
China Agricultural University  
sundx@cau.edu.cn  
https://orcid.org/0000-0002-7512-5946

Research

Keywords: Genome-wide-association study, immunoglobulins, SNP, Immune capacity, Chinese Holstein

DOI: https://doi.org/10.21203/rs.3.rs-40831/v1

License: ☑️ ☀️ This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background

Immunoglobulins (Igs) are important components of the innate immune system, and fight pathogens as a part of the first defense line. Newborn dairy calves get maternal antibodies from colostrum. Therefore, contents of immunoglobulins in colostrum and serum of cows are essential traits when estimating potential natural disease resistance of calves. In this study, a genome-wide association study (GWAS) was performed to identify candidate genes that are responsible for the observed genetic variation of immunoglobulins contents in colostrum and blood in Holstein cows.

Results

Colostrum, blood and hair follicle samples were collected from the 620 Chinese Holstein cows within 24 hours after calving. The concentration of IgG, IgG1, IgG2, IgA and IgM in both colostrum and serum were detected via ELISA methods, respectively. Using GCTA software, GWASs were performed with 88,934 SNPs genotyped by using Illumina 50K (54,609 SNPs) and GeneSeek 150K (140,668 SNPs) chips in which 50K chip were imputed to 150K SNPs with BEAGLE 3.0.4 software. As a result, 20 and 5 SNPs were detected genome-wide significantly associated with contents of the IgG and IgM in colostrum and serum (P<3.16E–6). In addition, 57, 11 and 10 SNPs were suggestive significantly associated with IgG, IgA and IgM traits (P<6.32E–5). Next, a total of 1,083 functional genes were identified that included or adjacent to these significant SNPs with a distance less than 1 Mb. Functional enrichment analysis showed that these genes were involved in immune related pathways, such as immune response, Fc gamma R-mediated phagocytosis, negative regulation of immunoglobulin secretion, humoral immune response, Fc-epsilon receptor and NF-kappaB signaling pathways. By integrating analysis of the functional enrichment and the known QTL data, we identified 21 candidate genes associated with contents of immunoglobulins in colostrum and serum, including ABR, TIMM22, CRK, MYO1C, RILP, SERPINF2, AKT1, BCL11B, HHIP1, DYNC1H1, HSP90AA1, TRAF3, KLC1, IL6, PYCARD, ITGAM, TGFB1I1, GUSB, CRCP, RABGEF1 and SBDS.

Conclusions

In this study, we identified 21 candidate genes associated with immunoglobulins level in colostrum and serum in dairy cattle. This founding demonstrated the possibility of increasing immunity through selective breeding and provided an important information for molecular breeding of dairy cattle.

Introduction

Maternal antibodies, especially immunoglobulin (Ig) G, are critical to early survival of domestic animals[1]. In dairy cattle, it is important for calves to gain maternal antibodies in the first 24 hours after born because the autoimmune system of newborn calves is too weak to resist various diseases.

According to the method for transmission of immunoglobulin from mother to young, mammals are divided into three categories, including only in utero (human and rabbit), only after birth (pig, horse, sheep and cow), or by both mechanisms (dog, rats and mice)[2]. Cow serum and lacteal secretions contain three major classes
of immunoglobulin antibodies: IgG, IgM and IgA. In a newborn calf, the IgG, IgM and IgA are absorbed from the colostrum into the circulation within 24-36 h after birth via a non-selective macromolecular transport system [3].

IgG molecule in dairy cattle occurs predominantly in two subclasses: IgG1 and IgG2. In the colostrum of cow which is defined as the secretion from the mammary gland during the first 24 h after calving, immunoglobulins make up 70–80% of the total protein content, whereas in mature milk, immunoglobulins account for only 1–2% of the protein [4, 5]. IgG1 comprises over 75% (46.4 mg/ml) of the immunoglobulin antibodies in colostral whey, followed by IgM (6.8 mg/ml), IgA (5.4 mg/ml) and IgG2 (2.9 mg/ml) [4]. Immunoglobulins are components of the innate immune system that can be found in animals without prior exposure to antigens, which have broad specificity [6]. Immunoglobulins participate principally the secondary immune response and possess a multitude of functions such as activate complement-mediated bacteriolytic reactions, augment the recognition and phagocytosis of bacteria by leucocytes (opsonization), prevent the adhesion of microbes to surfaces, inhibit bacterial metabolism, agglutinate bacteria, and neutralize toxins and viruses [7].

Heritability estimates for IgG concentrations in blood ranged from 0.27 to 0.64 in human [8-12]. In dairy cows, heritability estimates of immunoglobulins measured in blood or milk ranges from 0.08–0.45 [13-16], where IgM generally has the highest heritability estimates (0.18–0.45) and IgG has lower heritability estimates (0.08–0.31). Additionally, immunoglobulins level in serum had higher heritability than that in milk [13]. Previous GWASs have been performed for immunoglobulins content in serum or mature milk, the first study of these in 2,247 individuals from four European cohorts (CROATIA-Vis, CROATIA-Korcula, Orkney Complex Disease Study and Northern Swedish Population Health Study) identified 4 loci encoding glycosyltransferases associated with IgG N-glycans [17]. The GWAS for immunoglobulin G glycosylation patterns in human indicated that RUNX family transcription factor 3 (RUNX3) was associated with decreased galactosylation and involved in both IgA class switching and B-cell maturation as well as T-cell differentiation and apoptosis [18]. In pig, 4 significant SNPs were identified by GWAS performed for four traits, including interferon-gamma (IFN-c) and interleukin 10 (IL-10) levels, the ratio of IFN-c to IL-10 and IgG blocking percentage to CSFV in serum [19]. Especially, a previous GWAS for blood natural antibodies in Canadian Holstein cows found the 23 SNPs were significantly associated with IgG [12]. Another GWAS for milk natural antibodies in Dutch Holstein-Friesian cattle identified candidate genes on chromosome 17, 18, and 21 that related to immunoglobulin structure and early B cell development [20]. However, no study on immunoglobulin in colostrum have been reported in dairy cattle so far. In the present study, our objective is to identify candidate genes related to contents of immunoglobulins in colostrum and serum in dairy cattle and provide molecular information for genetically improving calves’ disease resistance.

Materials And Methods

Animals and phenotypes

The animals used in this study consists of 620 Chinese Holstein cows, the daughters of 44 sires with average 13.78 cows per sire. All the cows were from 10 dairy farms in the Beijing Dairy Cattle Center and the Beijing Sanyuan Lvhe Dairy Farming Center. The blood serum and colostrum samples were taken from each cow.
during the first milking within 24 h after calving for measurement of immunoglobulins. Also, hair follicle samples were collected from each animal for SNP chip genotyping. The cows were on average parity of 2.52, ranging from 1 to 4. The whole procedure for collection of the samples (blood, hair and colostrum) was carried out in strict accordance with the protocol approved by the Animal Welfare Committee of China Agricultural University (Permit number: DK996).

The commercially available, ELISA kits were used to measure the antibodies levels of each colostrum sample and serum sample according to the manufacturer’s instructions, including IgG (Bovine IgG ELISA Quantitation Set E10-118, Bethyl Laboratories, Montgomery, TX, USA), IgG1 (Bovine IgG1 ELISA Quantitation Set, E10-116), IgG2 (Bovine IgG2 ELISA Quantitation Set, E10-117), IgA (Bovine IgA ELISA Quantitation Set, E10-131) and IgM (Bovine IgM ELISA Quantitation Set, E10-101). Then, the phenotype values for the concentration of IgG, IgG1, IgG2 in colostrum or serum were square root transformed to accomplish normality, while phenotype values for IgA and IgM traits were log-transformed corrected. Thus, the adjusted phenotypic values were used for further statistical analysis.

Genotyping

Genomic DNA was extracted from hair follicles of each cow by using the QIAamp® DNA Mini Kit (QIAGEN, Valencia, CA, USA). Of all the cows, 588 individuals were implemented with the 150k chip (including 140,668 SNPs, GeneSeek, Lincoln, NE, USA) and 32 cows were genotyped with Illumina Bovine SNP50 BeadChip (including 54,609 SNPs, Illumina, San Diego, CA, USA).

Data imputation and quality control

To make most use of SNPs from high density chip, the total 32 samples genotyped with 50k chip were imputed to 150k chip based on our previously developed reference population that included 1198 Chinese Holstein cows by using the BEAGLE 3.0.4 software [21]. Allelic \( R^2 \) was estimated as an indicator of imputation accuracy based on the genotype probabilities.

Then we performed PLINK [22] for SNPs quality control and removed the SNPs with call rates < 90%, minor allele frequencies < 0.1, a deviation from Hardy-Weinberg equilibrium (HWE) \( P \) values < \( 10^{-6} \) and > 10 % missing genotypes [23]. After considering all the quality control measures, totally 88,934 SNPs were included in the genome-wide association study (Additional files 1 and 2). All SNP positions were determined according to the Bos taurus UMD 3.1 genome assembly.

Statistical analysis

Mixed model based single locus regression analyses (MMRA)

The association of the individual phenotype with each individual SNP was performed via following a mixed linear model:

\[
y = X\beta + g + A_g \sigma_g^2 + \sigma_e^2 \]
Where y is an n × 1 the vector of root square or log transformed corrected concentration immunoglobulins of all cow after calving with n being the sample size of 619; β is the vector of fixed effects, including herd, parity and season of calving, X is an incidence matrix relating elements of β to y; g is an n × 1 vector of the total genetic effects of the individuals with \( g \sim N(0, A\sigma_g^2) \), and A is interpreted as the genetic relationship matrix (GRM) between individuals. We can therefore estimate \( \sigma_g^2 \) by the restricted maximum likelihood (REML) approach [24], relying on the GRM estimated from all the SNPs. I is an n × 1 identity matrix, and \( \epsilon \) is a vector of residual effects with \( \epsilon \sim N(0, I\sigma_\epsilon^2) \).

GWAS was carried out by GCTA 1.90.2 software [25]. GCTA is the method which can estimate variance explained by all SNPs, and extend the method to partition the genetic variance onto each of the chromosomes and also to estimate the variance explained by the X chromosome.

In order to define the thresholds for genome-wide significant/suggestive associations, we calculated the effectively independent tests based on the estimated number of independent markers and linkage disequilibrium blocks for markers [26]. A linkage disequilibrium block was defined as a set of adjacent SNPs with pairwise \( r^2 \) values greater than 0.40 [27].

A total of 15,820 effectively independent tests was recommended and the threshold P-value for genome-wide significance association was 3.16E–6 (0.05/15,820) and for suggestive association was 6.32E–5 (1/15,820) suggested by Lander & Kruglyak [26, 28]. Inflation or deflation in p-values due to stratification or family structure was assessed by genomic inflation factor (\( \lambda \)) and also visually inspected by quantile-quantile (Q-Q) plot. \( \lambda \) is calculated as the median of the \( \chi^2 \) test statistics (1 degree of freedom) divided by its theoretical median under the null distribution[29]. The \( \lambda \) was calculated using GENABEL packages of R.

**Haplotype block analysis**

To further detect candidate regions associated with the immunoglobulins contents in colostrum and serum, we performed linkage disequilibrium analysis for the chromosomal regions with adjacent multiple significant SNPs based on Haploview v4.2[30]. The block was defined according to the solid spin algorithm by the criteria of Gabriel [31].

**Identification of candidate genes and functional annotation**

To identify the positional candidate genes that are potentially associated with the levels of immunoglobulins in colostrum and serum, the genomic regions that were close to the significant SNPs with a distance less than 1 Mb were selected. These regions were then referenced against the bovine genome assembly (UMD 3.1, Ensemble Genes Release 92) using the BioMart tool (http://asia.ensembl.org/biomart/martview) to found genes that are located in the vicinity of the significant SNPs. Simultaneously, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for these extracted genes by KOBAS 3.0 [32]. KOBAS annotates a set of genes with putative pathways and immunity relationships by mapping to genes with a known annotation. Afterwards, we compared the physical positions of the identified candidate genes with the reported quantitative traits loci (QTLs) that have been shown to be associated with immune capacity traits for cattle in the Animal QTL database (http://www.animalgenome.org/cgi-bin/QTLdb/index).
Results

Statistics of phenotype and SNP data

In this study, we analyzed 10 traits, including concentration of IgG, IgG1, IgG2, IgA and IgM in colostrum and serum. Mean and the corresponding standard deviations for the original and corrected phenotypic values with or without square root or log transformation were shown in Table 1.

Genome-wide association study

For the genotype imputation with the software Beagle 3.0.4, SNPs with \( R^2 > 0.3 \) were retained [33] and the average allelic \( R^2 \) of imputed genotypes was 74.8%, which made sure the accuracy of imputation and full use of chip data. With GCTA 1.90.2, no population substructure was observed from QQ plots (Figures. 1). In addition, the inflation factor (\( \lambda \)) estimated was from 0.98-1.02 for all traits, which indicated that our results can be accepted for further analysis.

For the contents of immunoglobulins in colostrum, as shown in Figure. 2 and Table 2, at genome-wide significant level, 3 SNPs (P<3.16E-6) located on BTA-20 (1 SNP) and 21 (2 SNPs) were found for concentration of IgG2 in colostrum. At suggestive significant level, a total of 11, 7 and 17 suggestive significant SNPs (P<6.32E-5) were detected for concentration of IgG, IgG1 and IgG2 in colostrum. Of these, the majority of them were located on BTA-21 (15 SNPs) while a few SNPs were located on BTA-1 (3 SNPs), 23 (2 SNPs), 24 (7 SNPs), 3 (2 SNPs), 5 (1 SNP), 6 (1 SNP), 20 (2 SNPs), and 27 (2 SNPs). In addition, a total of 3 and 6 suggestive significant SNPs (P<6.32E-5) were detected for the concentration of IgA and IgM in colostrum and distributed in chromosomes 1 (1 SNP), 4 (1 SNP), 5 (1 SNP), 8 (1 SNP), 10 (2 SNPs), 15 (1 SNP) and 17 (2 SNPs). Associations between 4 SNPs and concentration of IgM in colostrum reached genome-wide significant level in chromosome 10 (2 SNPs) and 17 (2 SNPs) (P<3.16E-6).

For the immunoglobulins contents in serum, genome-wide significant associations between 17 SNPs and concentration of IgG2 were found in BTA-21 (16 SNPs) and 20 (1 SNP) (P<3.16E-6). Additionally, 22 suggestive significant SNPs (P<6.32E-5) were detected for IgG1 and IgG2 in serum. Of these, 15 SNPs were located on BTA-21 and the rest of them were on 28 (1 SNP), 19 (1 SNP), 7 (1 SNP), 19 (1 SNP), 12 (2 SNPs), and 20 (1 SNP) (Figure. 2 and Table 3). At both level, concentration of IgG in serum did not show any significant associations. For IgA and IgM in serum, 8 and 4 suggestive significant SNPs (P<6.32E-5) were identified in chromosomes 2 (1 SNP), 15 (3 SNPs), 18 (1 SNP), 25 (2 SNP), 29 (1 SNP), 4 (1 SNP), 6 (1 SNP) and 7 (2 SNPs) (Figure. 2 and Table 3). And 1 SNP associate with IgA in BTA-17 achieved genome-wide significant level (P<3.16E-6). The Manhattan plots for concentrations of IgG, IgG1, IgG2 IgA and IgM in colostrum and serum were shown in Figures. 2.

Haplotype Block Analysis

As shown in Figure. 2, with regards to the concentration of IgG in colostrum, 7 significant SNPs on BAT 24 (45.89–46.24Mb) generated a haplotype block (Figure. 3A) and 22 genes were found within this region after gene searching. As for the concentration of IgG2 in colostrum, 18 significant SNPs were found located on BTA-21 (66.58–71.39Mb) that contained 168 genes. In this region, we detected 3 haplotype blocks: block
1 (14kb, 2 SNPs), block 2 (17kb, 3 SNPs) and block 3 (14kb, 2 SNPs) (Figure 3B). For the concentration of IgG in serum, we detected 32 significant SNPs on BAT 21 (64.93–72.48Mb) that accommodate 204 genes and totally formed 3 haplotype blocks, including block 1 (18kb, 2 SNPs), block 2 (62kb, 4 SNPs) and block 3 (14kb, 2 SNPs) (Figure 3C). Additionally, the haplotype block of IgG concentration in colostrum was covered in the block of IgG concentration in serum. However, few significant SNPs were observed within haplotype blocks for the other traits.

Candidate Genes and Function Analysis

For IgG in serum and colostrum, a total of 707 functional genes that contained or were near (within 1Mb) the identified significant SNPs were obtained based on the bovine genome assembly UMD3.1. In addition, 252 and 408 functional genes were identified that contained or were near the identified significant SNPs for IgA and IgM traits with distance of less than 1 Mb, respectively. After removing the duplicates of results, a total of 1,083 candidate genes were obtained for contents of immunoglobulins traits, 706 protein-coding genes, 60 miRNA genes, 45 spliceosomal RNAs, 76 small nucleolar RNAs and 196 novel genes (Additional file 3: Table S1). Out of these, with GO and KEGG analysis, 151 genes were observed participated in immune related pathways such as immune response, Fc gamma R-mediated phagocytosis, negative regulation of immunoglobulin secretion, humoral immune response, Fc-epsilon receptor and NF-kappaB signaling pathways, etc (Additional file 4: Table S2).

Based on the Cattle QTL database that has released 232 loci for immune capacity until now (April 26, 2020, http://www.animalgenome.org/cgi-bin/QTLdb/), we compared the physical positions of the 151 candidate genes with the peak of the known QTLs for immune capacity in dairy cattle, including IgG level, FMDV peptide-induced cell proliferation and ConA-induced cell proliferation. Consequently, 21 genes were found located within the QTL regions with distance to the peak positions of less than 1.0 cM so that they were considered as promising candidates for immunoglobulins level in serum and colostrum (Table 4). They were BR activator of RhoGEF and GTPase (ABR), translocase of inner mitochondrial membrane 22 (TIMM22), CRK proto-oncogene, adaptor protein (CRK), myosin IC (MYO1C), Rab interacting lysosomal protein (RILP), serpin family F member 2 (SERPINF2), AKT serine/threonine kinase 1 (AKT1), BAF chromatin remodeling complex subunit BCL11B (BCL11B), HHIPL like 1 (HHIPL1), dynein cytoplasmic 1 heavy chain 1 (DYNC1H1), heat shock protein 90 alpha family class A member 1 (HSP90AA1), TNF receptor associated factor 3 (TRAF3), kinesin light chain 1 (KLC1), interleukin 6 (IL6), PYD and CARD domain containing (PYCARD), integrin subunit alpha M (ITGAM), transforming growth factor beta 1 induced transcript 1 (TGFB1I1), glucuronidase beta (GUSB), CGRP receptor component (CRCP), RAB guanine nucleotide exchange factor 1 (RABGEF1) and SBDS ribosome maturation factor (SBDS).

Discussion

In this study, we identified candidate region influencing immunoglobulins in colostrum and serum by performing GWASs with high density SNP genotypes. We found 3 significant regions on 21, 10 and 17 chromosome and 78 suggestive significant SNPs on other 20 chromosomes associated with IgG, IgA and IgM. To our knowledge, this is the first investigation focusing on unrevealing the genetic mechanism of these immune traits in bovine colostrum based on a high-density SNP chip panel.
Population stratification is a major factor for false positives in GWAS. In general, a genomic inflation factor $\lambda$ of $<1.05$ indicates no population stratification [34]. In this study, $\lambda$ values were 0.98~1.02 for concentration and ratio of IgG, IgM and IgA in colostrum and serum. QQ plots indicated that population stratification has been controlled as well. In addition, we calculated the effectively independent tests based on the estimated number of independent markers and linkage disequilibrium blocks for autosome markers, and a total of 15,820 effectively independent tests was suggested. This setting of thresholds may lead slight false positive but numerous useful information.

In this GWAS, we found a region on BTA-21 from 63.9 to 71.5 Mb that was significantly associated with for IgG2. Previous GWASs on bovine nature antibodies in Canadian and Dutch Holstein populations identified genomic regions on BTA-21 from 55.5 to 70.6 Mb and 66.0 to 71.6 Mb that were significantly associated with KLH-IgG and PGN-IgG1 [12, 20], which overlapped with region in this study. Although our samples were taken from colostrum and serum, whereas theirs were from serum and mature milk, and we analyzed IgG, IgG1 and IgG2 while they analyzed IgG and IgG1, it is likely due to the high genetic correlation between the same IgG isotype in serum and milk (0.81±0.18) [13]. This implies that the all IgG isotype subclasses from serum or milk shared the same genetic mechanism.

For IgM, although its heritability was relatively high, we detected 10 and 5 SNPs associate with concentration of IgM in colostrum and serum, respectively, we did not identified any candidate genes for this trait. Also, the identified significant SNPs for IgM in colostrum and serum were totally different. Similar results were reported in a previous GWAS by de Klerk et al. [12]. As for the other genomic regions we identified for immunoglobulins have not been reported in previous studies in bovine.

**Role of Candidate Genes**

For IgG1 in colostrum and serum, 6 candidate genes were identified and from 22.3~23.4 Mb on BTA-19. *ABR*, encodes protein that contains a GTPase-activating protein domain. Previous studies identified *ABR* and *BCR* as the only GTPase-activating proteins that specifically negatively regulate Rac function in vivo in primary macrophages and normally curb very specific functions of mature tissue innate immune cells [35, 36]. The protein encoded by *TIMM22* is a subunit of *TIM22* which is transporters and represents the voltage-activated and signal-gated channel. *TIMM22* is involved in protein transmembrane transporter activity, which may relate to the transport of IgG though the receptor. *CRK* encodes a member of an adapter protein family that binds to several tyrosine-phosphorylated proteins. Differential migration of CRK/CRKL-deficient T cells resulted in efficient graft-versus-leukemia responses in mice [37]. *MYO1C* encodes a member of the unconventional myosin protein family which play a role in bacterial infectious disease and Pathogenic Escherichia coli infection. *RILP* encodes a lysosomal protein that is related to the phagosomes in macrophages [38]. *SERPINF2* encodes a member of the serpin family of serine protease inhibitors. The protein is a major inhibitor of plasmin, which degrades fibrin and various other proteins. *SERPINF2* is involved in the recruitment of lymphocytes in the peripheral tissues and the progression of fibrosis. Besides, *SERPINF2* contributes to the maintenance of immunological functions that are related to IgE in human [39].

For IgG2, totally 7 candidate genes were found on BTA-21 from 65.8~70.9 Mb. *AKT1*, is part of the PI3K-AKT pathway and involved in B cell receptor, T cell receptor, Fc gamma R-mediated phagocytosis pathway. Some
of the known functions include T cell development and FOXO1 signaling, which involves maturation and survival of peripheral B cells and class switching [40-42]. It is worth mentioning that AKT1 was also identified as candidate gene for IgG1 in a previous study [20]. BCL11B encodes a C2H2-type zinc finger protein that regulates the initial stages of human T-cell differentiation and mediated epigenetic repression for histone deacetylase inhibitors in cutaneous T-Cell lymphoma [43-45]. HHIPL1 belongs to the glucose/sorbosone dehydrogenase family and take part in receptor-mediated endocytosis so that it can effect receptor activity. DYNC1H1 encodes a member of the cytoplasmic dynein heavy chain family that function as molecular motors and play a role in Salmonella infection. The protein encoded by HSP90AA1 is an inducible molecular chaperone that functions as a homodimer. Hsp90 was related to beta cell autoimmunity and protected cells from complement-dependent cytotoxicity by inhibiting, together with mortalin, C5b-9 assembly and/or stability at the plasma membrane [46, 47]. The protein encoded by TRAF3 is a member of the TNF receptor associated factor (TRAF) protein family that participates in the signal transduction of CD40, a TNFR family member important for the activation of the immune response and was found to be a critical component of the lymphotoxin-beta receptor (LTbetaR) signaling complex, which induces NF-kappaB activation and cell death initiated by LTbeta ligation [48, 49]. The findings suggest that Parkin plays a novel role in innate immune signaling by targeting TRAF3 for degradation and maintaining the balance of innate antiviral immunity [50]. A TRAF3-NIK axis differentially regulates viral DNA vs RNA pathways in innate immune signaling [51]. KLC1 encodes a member of the kinesin light chain family. KLC1 is associated with infectious disease caused by bacterial. Additionally, 2 and 4 significant SNPs for the ratio of IgG1 and IgG2 in colostrum and serum but no candidate gene were selected, which might attribute to the less transmission but most self-synthetization.

For IgA, we selected 7 candidate genes. PYCARD, encodes an adaptor protein that is composed of two protein-protein interaction domains: a N-terminal PYRIN-PAAD-DAPIN domain (PYD) and a C-terminal caspase-recruitment domain (CARD). The PYD and CARD domains are members of the six-helix bundle death domain-fold superfamily that mediates assembly of large signaling complexes in the inflammatory and apoptotic signaling pathways via the activation of caspase. In normal cells, this protein is localized to the cytoplasm. Han Y et.al found that PYCARD as a negative regulator of the MAVS-mediated innate immunity, may play an important role in host protection upon virus infection [52]. ITGAM gene encodes the integrin alpha M chain. The alpha M beta 2 integrin is important in the adherence of neutrophils and monocytes to stimulated endothelium, and also in the phagocytosis of complement coated particles and ITGAM as a ligand for the integrin Mac-1 and suggest that many immune-modulating effects previously ascribed to PF4 are mediated through its interaction with Mac-1 [53]. TGFB1I1 encodes a coactivator of the androgen receptor, a transcription factor which is activated by androgen and has a key role in male sexual differentiation. Hic-5 regulates GR binding site selection by a novel mechanism, exploiting gene-specific requirements for chromatin remodeling enzymes to selectively influence DNA occupancy and gene regulation by a transcription factor [54]. GUSB encodes a hydrolase that degrades glycosaminoglycans, including heparan sulfate, dermatan sulfate, and chondroitin-4,6-sulfate and play role in HCV-infected human liver [55]. CRCP encodes a membrane protein that functions as part of a receptor complex for a small neuropeptide that increases intracellular cAMP levels. The influence of the CGRP family peptides in reproduction and pregnancy is explored and discussed[56]. RABGEF1 forms a complex with rabaptin-5 (RABPT5; MIM 603616) that is required for endocytic membrane fusion. Millrine D et al. identified Rabex-5 as an immunomodulatory drugs (IMiDs) target molecule that functions to restrain TLR activated auto-immune promoting pathways. We
propose that release of Rabex-5 from complex with Cereblon enables the suppression of immune responses, contributing to the anti-inflammatory properties of IMiDs [57]. Additionally, RABGEF1 mediates recycling endosome fusion with GAS-containing autophagosome-like vacuoles through the STX6-VAMP3-VTI1B complex; SNAREs are involved in autophagosome formation in response to bacterial infection [58]. SBDS encodes a highly conserved protein that related to primary immunodeficiency and phagocyte defects. In addition, SBDS play a role in immune system process and hematopoietic or lymphoid organ development. Another candidate gene in chromosome 4, IL6, encodes a cytokine that functions in inflammation and the maturation of B cells. Besides, IL6 has been shown to have important functions in immune response, hematopoiesis, inflammation and the acute phase response. The protein is primarily produced at sites of acute and chronic inflammation, where it is secreted into the serum and induces a transcriptional inflammatory response through interleukin 6 receptor, alpha [59, 60]. In the breeding perspective of dairy cattle, the relatively high heritability of immunoglobulin and available genetic information provides an opportunity to include immunoglobulin as a new trait to improve dairy cattle health. Although plentiful studies for the benefits of IgM and to a lesser extent IgG in mice [61-63], the beneficial correlations reported between immunoglobulin and resistance against infectious diseases are minimal in livestock [15, 64, 65]. Hence, the character of immunoglobulin in immunity of dairy cattle needs further verification. It cannot be ignored that a number of studies are proposing self-antigens as the stimulator of B1 cells and reporting both negative and positive roles of self-reactive immunoglobulin in autoimmune diseases [62]. In mice, lupus-like autoimmune diseases developed when immunoglobulin production was impaired, but mice that could produce only IgM immunoglobulin, but not IgG immunoglobulin, did not develop the autoimmune disease [66, 67]. Therefore, to develop breeding strategies to increase immunoglobulin in animals with relatively long lifespan such as cattle the antigen should be cautiously explored and animals monitored for any signs of autoimmunity.

Conclusion

Natural antibodies immunoglobulins in colostrum and serum have an important genetic component, and we identified specific variants in the genome that influence immunoglobulins by GWAS. A total of 25 genome wide and 78 suggestive significant SNPs were identified for IgG, IgG1, IgG2, IgA and IgM in both colostrum and serum. In the case of IgG, chromosomes-19, 21, carried the main quantitative trait loci. Candidate regions for IgA were main located in chromosomes-4 and 25. Integrated analysis of the identified significant SNPs, functional enrichment and the known QTL data, we suggested 21 candidate genes for content of immunoglobulins in colostrum and serum. Of these, ABR, TIMM22, CRK, MYO1C, RILP, SERPINF2, AKT1, BCL11B, HHIP1L1, DYNC1H1, HSP90AA1, TRAF3 and KLC1 were for IgG traits, as well as IL6, PYCARD, ITGAM, TGFB1I1, GUSB, CRCP, RABGEF1, SBDS were identified candidates for IgA content in colostrum and serum, respectively. Our results provide new insights into the regulation of immunoglobulins that will help give a better understanding of the complex relationship between milk nature antibodies and disease resistance in dairy cattle and provide critical genetic information for management or genome selection.

Declarations

Acknowledgements
We appreciate Beijing Dairy Cattle Center for providing the animal samples.

**Funding**

This work was financially supported by the National Natural Science Foundation of China (31872330, 31872330, 31802041), Beijing Science and Technology Program (20200105, D171100002417001), and the Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62).

**Availability of data and materials**

All supporting data can be found within the additional files.

**Authors’ contributions**

SL performed bioinformatics and statistical analysis, and also was a major contributor to manuscript preparation. CK, LL, YG and XL performed experiments and sample collection. CK and BH participated in result interpretation, wrote, revised and approved the manuscript. YZ and SZ commented the manuscript. DS conceived and designed the experiments and was the major contribution to manuscript revision. All authors read and approved the final manuscript.

**Ethics approval and consent to participate**

All protocols for collection of the blood and hair samples of China Holstein cows were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at China Agricultural University. Blood and hair samples were collected specifically for this study following standard procedures with the full agreement of the Beijing Dairy Cattle Center who owned the animals.

**Consent for Publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**References**

1. Cabrera RA, Lin X, Campbell JM, Moeser AJ, Odle J: Influence of birth order, birth weight, colostrum and serum immunoglobulin G on neonatal piglet survival. *J Anim Sci Biotechnol* 2012, 3(1):42.
2. Butler JE, Wertz N, Sinkora M: Antibody Repertoire Development in Swine. *Annu Rev Anim Biosci* 2017, 5:255-279.
3. McFadden TB, Besser TE, Barrington GM: Regulation of immunoglobulin transfer into mammary secretions of ruminants. *Biotechn Ag* 1997(18):133-152.
4. Porter P: Immunoglobulins in bovine mammary secretions. Quantitative changes in early lactation and absorption by the neonatal calf. *Immunology* 1972, 23(2):225-238.
5. Leach JL, Sedmak DD, Osborne JM, Rahill B, Lairmore MD, Anderson CL: Isolation from human placenta of the IgG transporter, FcRn, and localization to the syncytiotrophoblast - Implications for maternal-fetal antibody transport. *Journal of Immunology* 1996, 157(8):3317-3322.

6. Baumgarth N, Tung JW, Herzenberg LA: Inherent specificities in natural antibodies: a key to immune defense against pathogen invasion. *Springer Semin Immun* 2005, 26(4):347-362.

7. Korhonen H, Marnila P, Gill HS: Milk immunoglobulins and complement factors. *Br J Nutr* 2000, 84 Suppl 1:S75-80.

8. Grundbacher FJ: Heritability estimates and genetic and environmental correlations for the human immunoglobulins G, M, and A. *Am J Hum Genet* 1974, 26(1):1-12.

9. Kacprazak-Bergman I: Sexual dimorphism of heritability of immunoglobulin levels. *Ann Hum Biol* 1994, 21(6):563-569.

10. Outschoorn IM, Talor MV, Hoffman WH, Rowley MJ, Mackay IR, Rose NR, Burek CL: Autoantibody heritability in thyroiditis: IgG subclass contributions. *Autoimmunity* 2011, 44(3):195-200.

11. Outschoorn IM, Talor MV, Burek CL, Hoffman WH, Rose NR: Heritability analysis of IgG4 antibodies in autoimmune thyroid disease. *Autoimmunity* 2014, 47(5):320-326.

12. de Klerk B, Emam M, Thompson-Crispi KA, Sargolzaei M, van der Poel JJ, Mallard BA: A genome-wide association study for natural antibodies measured in blood of Canadian Holstein cows. *BMC Genomics* 2018, 19(1):694.

13. de Klerk B, Ducro BJ, Heuven HC, den Uyl I, van Arendonk JA, Parmentier HK, van der Poel JJ: Phenotypic and genetic relationships of bovine natural antibodies binding keyhole limpet hemocyanin in plasma and milk. *J Dairy Sci* 2015, 98(4):2746-2752.

14. Ploegaert TC, Wijga S, Tijhaar E, van der Poel JJ, Lam TJ, Savelkoul HF, Parmentier HK, van Arendonk JA: Genetic variation of natural antibodies in milk of Dutch Holstein-Friesian cows. *J Dairy Sci* 2010, 93(11):5467-5473.

15. Thompson-Crispi KA, Miglior F, Mallard BA: Genetic parameters for natural antibodies and associations with specific antibody and mastitis in Canadian Holsteins. *J Dairy Sci* 2013, 96(6):3965-3972.

16. Wijga S, Bovenhuis H, Bastiaansen JWM, van Arendonk JAM, Ploegaert TCW, Tijhaar E, van der Poel JJ: Genetic parameters for natural antibody isotype titers in milk of Dutch Holstein-Friesians. *Animal Genetics* 2013, 44(5):485-492.

17. Lauc G, Huffman JE, Pucic M, Zgaga L, Adamczyk B, Muzinic A, Novokmet M, Polasek O, Gornik O, Kristic J et al.: Loci Associated with N-Glycosylation of Human Immunoglobulin G Show Pleiotropy with Autoimmune Diseases and Haematological Cancers. *Plos Genet* 2013, 9(1).

18. Wahl A, van den Akker E, Klaric L, Stambuk J, Benedetti E, Plomp R, Razdorov G, Trbojevic-Akmacic I, Deelen J, van Heemst D et al.: Genome-Wide Association Study on Immunoglobulin G Glycosylation Patterns. *Front Immunol* 2018, 9.

19. Lu X, Liu J, Fu W, Zhou J, Luo Y, Ding X, Liu Y, Zhang Q: Genome-wide association study for cytokines and immunoglobulin G in swine. *PLoS One* 2013, 8(10):e74846.

20. Cordero-Solorzano J, Parmentier HK, Arts JAJ, van der Poel J, de Koning DJ, Bovenhuis H: Genome-wide association study identifies loci influencing natural antibody titers in milk of Dutch Holstein-Friesian
cattle. *J Dairy Sci* 2019, 102(12):11092-11103.

21. Browning BL, Browning SR: A Unified Approach to Genotype Imputation and Haplotype-Phase Inference for Large Data Sets of Trios and Unrelated Individuals. *American Journal of Human Genetics* 2009, 84(2):210-223.

22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ *et al.*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007, 81(3):559-575.

23. Wiggans GR, Sonstegard TS, VanRaden PM, Matukumalli LK, Schnabel RD, Taylor JF, Schenkel FS, Van Tassell CP: Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *Journal of Dairy Science* 2009, 92(7):3431-3436.

24. Patterson HD, Thompson R: Recovery of Inter-Block Information When Block Sizes Are Unequal. *Biometrika* 1971, 58(3):545-+.

25. Yang J, Lee SH, Goddard ME, Visscher PM: GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011, 88(1):76-82.

26. Nicodemus KK, Liu WL, Chase GA, Tsai YY, Fallin MD: Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms. *Bmc Genet* 2005, 6.

27. Gu XR, Feng CG, Ma L, Song C, Wang YQ, Da Y, Li HF, Chen KW, Ye SH, Ge CR *et al.*: Genome-Wide Association Study of Body Weight in Chicken F2 Resource Population. *Plos One* 2011, 6(7).

28. Lander E, Kruglyak L: Genetic Dissection of Complex Traits - Guidelines for Interpreting and Reporting Linkage Results. *Nat Genet* 1995, 11(3):241-247.

29. Devlin B, Roeder K: Genomic control for association studies. *Biometrics* 1999, 55(4):997-1004.

30. Barrett JC, Fry B, Maller J, Daly MJ: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005, 21(2):263-265.

31. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M *et al.*: The structure of haplotype blocks in the human genome. *Science* 2002, 296(5576):2225-2229.

32. Xie C, Mao XZ, Huang JJ, Ding Y, Wu JM, Dong S, Kong L, Gao G, Li CY, Wei LP: KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 2011, 39:W316-W322.

33. Browning SR, Browning BL: Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 2007, 81(5):1084-1097.

34. Price AL, Zaitlen NA, Reich D, Patterson N: New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 2010, 11(7):459-463.

35. Cho YJ, Cunnick JM, Yi SJ, Kaartinen V, Groffen J, Heisterkamp N: Abr and Bcr, two homologous Rac GTPase-activating proteins, control multiple cellular functions of murine macrophages. *Mol Cell Biol* 2007, 27(3):899-911.
36. Cunnick JM, Schmidhuber S, Chen G, Yu M, Yi SJ, Cho YJ, Kaartinen V, Minoo P, Warburton D, Groffen J et al. Bcr and Abr cooperate in negatively regulating acute inflammatory responses. *Mol Cell Biol* 2009, 29(21):5742-5750.

37. Huang Y, Clarke F, Karimi M, Roy NH, Williamson EK, Okumura M, Mochizuki K, Chen EJ, Park TJ, Debes GF et al. CRK proteins selectively regulate T cell migration into inflamed tissues. *J Clin Invest* 2015, 125(3):1019-1032.

38. Seto S, Matsumoto S, Tsujimura K, Koide Y: Differential recruitment of CD63 and Rab7-interacting-lysosomal-protein to phagosomes containing Mycobacterium tuberculosis in macrophages. *Microbiol Immunol* 2010, 54(3):170-174.

39. Okada K, Ueshima S, Kawao N, Yano M, Tamura Y, Tanaka M, Sakamoto A, Hatano M, Arima M, Miyata S et al. Lack of both alpha2-antiplasmin and plasminogen activator inhibitor type-1 induces high IgE production. *Life Sci* 2013, 93(2-3):89-95.

40. Juntilla MM, Wofford JA, Birnbaum MJ, Rathmell JC, Koretzky GA: Akt1 and Akt2 are required for alphabeta thymocyte survival and differentiation. *Proc Natl Acad Sci U S A* 2007, 104(29):12105-12110.

41. Calamito M, Juntilla MM, Thomas M, Northrup DL, Rathmell J, Birnbaum MJ, Koretzky G, Allman D: Akt1 and Akt2 promote peripheral B-cell maturation and survival. *Blood* 2010, 115(20):4043-4050.

42. Chen J, Limon JJ, Blanc C, Peng SL, Fruman DA: Foxo1 regulates marginal zone B-cell development. *Eur J Immunol* 2010, 40(7):1890-1896.

43. Ha VL, Luong A, Li F, Casero D, Malvar J, Kim YM, Bhatia R, Crooks GM, Parekh C: The T-ALL related gene BCL11B regulates the initial stages of human T-cell differentiation. *Leukemia* 2017, 31(11):2503-2514.

44. Fu WJ, Yi SG, Qiu L, Sun JR, Tu P, Wang Y: BCL11B-Mediated Epigenetic Repression Is a Crucial Target for Histone Deacetylase Inhibitors in Cutaneous T-Cell Lymphoma. *J Invest Dermatol* 2017, 137(7):1523-1532.

45. Grabarczyk P, Winkler P, Delin M, Sappa PK, Bekeschus S, Hildebrandt P, Przybylski GK, Volker U, Hammer E, Schmidt CA: The N-Terminal CCHC Zinc Finger Motif Mediates Homodimerization of Transcription Factor BCL11B. *Mol Cell Biol* 2018, 38(5).

46. Nakai A, Minematsu T, Tamai N, Sugama J, Urai T, Sanada H: Prediction of healing in Category I pressure ulcers by skin blotting with plasminogen activator inhibitor 1, interleukin-1 alpha, vascular endothelial growth factor C, and heat shock protein 90 alpha: A pilot study. *J Tissue Viability* 2019, 28(2):87-93.

47. Ocana GJ, Sims EK, Watkins RA, Ragg S, Mather KJ, Oram RA, Mirmira RG, DiMeglio LA, Blum JS, Evans-Molina C: Analysis of serum Hsp90 as a potential biomarker of beta cell autoimmunity in type 1 diabetes. *Plos One* 2019, 14(1).

48. Zhang JL, Chen T, Yang XP, Cheng H, Spath SS, Clavijo PE, Chen JH, Silvin C, Issaeva N, Su XL et al: Attenuated TRAF3 Fosters Activation of Alternative NF-kappa B and Reduced Expression of Antiviral Interferon, TP53, and RB to Promote HPV-Positive Head and Neck Cancers. *Cancer Res* 2018, 78(16):4613-4626.

49. Bishop GA, Stunz LL, Hostager BS: TRAF3 as a Multifaceted Regulator of B Lymphocyte Survival and Activation. *Front Immunol* 2018, 9.
50. Xin D, Gu HY, Liu EP, Sun QM: Parkin negatively regulates the antiviral signaling pathway by targeting TRAF3 for degradation. *J Biol Chem* 2018, 293(31):11996-12010.

51. Parvatiyar K, Pindado J, Dev A, Aliyari SR, Zaver SA, Gerami H, Chapon M, Ghaffari AA, Dhingra A, Cheng GH: A TRAF3-NIK module differentially regulates DNA vs RNA pathways in innate immune signaling. *Nature Communications* 2018, 9.

52. Han Y, He X, Wei C, Song T, Zou L, Li Z, Ye J, Qi L, Li L, Zhong H *et al*.: Negative regulation of MAVS-mediated innate immune response by ASC. *Mol Cell Biochem* 2018, 445(1-2):35-43.

53. Lishko VK, Yakubenko VP, Ugarova TP, Podolnikova NP: Leukocyte integrin Mac-1 (CD11b/CD18, alphaMbeta2, CR3) acts as a functional receptor for platelet factor 4. *J Biol Chem* 2018, 293(18):6869-6882.

54. Chan LP, Liu C, Chiang FY, Wang LF, Lee KW, Chen WT, Kuo PL, Liang CH: IL-8 promotes inflammatory mediators and stimulates activation of p38 MAPK/ERK-NF-kappaB pathway and reduction of JNK in HNSCC. *Oncotarget* 2017, 8(34):56375-56388.

55. Congiu M, Slavin JL, Desmond PV: Expression of common housekeeping genes is affected by disease in human hepatitis C virus-infected liver. *Liver Int* 2011, 31(3):386-390.

56. Yallampalli C, Chauhan M, Endsley J, Sathishkumar K: Calcitonin gene related family peptides: importance in normal placental and fetal development. *Adv Exp Med Biol* 2014, 814:229-240.

57. Millrine D, Tei M, Gemechu Y, Kishimoto T: Rabex-5 is a lenalidomide target molecule that negatively regulates TLR-induced type 1 IFN production. *Proc Natl Acad Sci U S A* 2016, 113(38):10625-10630.

58. Nozawa T, Minowa-Nozawa A, Aikawa C, Nakagawa I: The STX6-VTI1B-VAMP3 complex facilitates xenophagy by regulating the fusion between recycling endosomes and autophagosomes. *Autophagy* 2017, 13(1):57-69.

59. Zhao S, Zhong Y, Fu X, Wang Y, Ye P, Cai J, Liu Y, Sun J, Mei Z, Jiang Y *et al*.: H3K4 Methylation Regulates LPS-Induced Proinflammatory Cytokine Expression and Release in Macrophages. *Shock* 2019, 51(3):401-406.

60. Frampah B, Luckett-Chastain LR, Gallucci RM: IL-6 Negatively Regulates IL-22Ralpha Expression on Epidermal Keratinocytes: Implications for Irritant Contact Dermatitis. *J Immunol Res* 2019, 2019:6276254.

61. Panda S, Ding JL: Natural antibodies bridge innate and adaptive immunity. *J Immunol* 2015, 194(1):13-20.

62. Mannoor K, Xu Y, Chen C: Natural autoantibodies and associated B cells in immunity and autoimmunity. *Autoimmunity* 2013, 46(2):138-147.

63. Gronwall C, Silverman GJ: Natural IgM: beneficial autoantibodies for the control of inflammatory and autoimmune disease. *J Clin Immunol* 2014, 34 Suppl 1:S12-21.

64. van Altena SE, Peen MA, van der Linden FH, Parmentier HK, Savelkoul HF, Tijhaar EJ: Bovine natural antibodies in antibody-dependent bactericidal activity against Escherichia coli and Salmonella Typhimurium and risk of mastitis. *Vet Immunol Immunopathol* 2016, 171:21-27.

65. Berghof TVL, Arts JAJ, Bovenhuis H, Lammers A, van der Poel JJ, Parmentier HK: Antigen-dependent effects of divergent selective breeding based on natural antibodies on specific humoral immune...
responses in chickens. *Vaccine* 2018, 36(11):1444-1452.

66. Jiang C, Foley J, Clayton N, Kissling G, Jokinen M, Herbert R, Diaz M: Abrogation of lupus nephritis in activation-induced deaminase-deficient MRL/lpr mice. *J Immunol* 2007, 178(11):7422-7431.

67. Jiang C, Zhao ML, Scearce RM, Diaz M: Activation-induced deaminase-deficient MRL/lpr mice secrete high levels of protective antibodies against lupus nephritis. *Arthritis Rheum* 2011, 63(4):1086-1096.

**Tables**

**Table 1:** Means and standard deviations for the original and corrected concentrations of immunoglobulins; n = 620

| Traits   | Mean | Std dev | Min | Max  | Traits   | Mean | Std dev | Min | Max  |
|----------|------|---------|-----|------|----------|------|---------|-----|------|
| col_IgG  | 47.44| 29.02   | 0.01| 146.60| col_IgG_sqrt| 6.51 | 2.23    | 0.10| 12.11|
| col_IgG1 | 14.46| 7.75    | 0.01| 37.73| col_IgG1_sqrt| 3.63 | 1.14    | 0.05| 6.14 |
| col_IgG2 | 3.20 | 1.94    | 0.07| 11.04| col_IgG2_sqrt| 1.70 | 0.54    | 0.26| 3.32 |
| ser_IgG  | 8.34 | 3.37    | 0.68| 22.90| ser_IgG_sqrt| 2.83 | 0.58    | 0.82| 4.79 |
| ser_IgG1 | 1.04 | 0.74    | 0.01| 7.65 | ser_IgG1_sqrt| 1.04 | 0.31    | 0.09| 2.77 |
| ser_IgG2 | 13.79| 6.51    | 0.69| 42.36| ser_IgG2_sqrt| 3.61 | 0.86    | 0.83| 6.51 |
| col_IgA  | 3.11 | 3.59    | 0.01| 51.18| col_IgA_log | 0.28 | 0.49    | -1.91| 1.71 |
| col_IgM  | 5.47 | 3.70    | 0.09| 36.41| col_IgM_log | 0.63 | 0.35    | -1.06| 1.56 |
| ser_IgA  | 0.23 | 0.17    | 0.01| 1.21 | ser_IgA_log | -0.73| 0.39    | -5.40| 0.08 |
| ser_IgM  | 2.25 | 1.68    | 0.01| 14.93| ser_IgM_log | 0.24 | 0.34    | -2.34| 1.17 |

Mean, arithmetic mean; SD, standard deviation; Max, maximum; Min, minimum; col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represent the corresponding concentration value of IgG, IgG1 IgG2, IgA and IgM in colostrum, ser_IgG, ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represent the corresponding concentration value of IgG, IgG1 IgG2, IgA and IgM in serum, respectively; _sqrt, phenotype by square root transformed; _log, phenotype by log transformed.
Table 2: The significant SNPs for concentrations of IgG, IgG1, IgG2, IgA and IgM in colostrum.

| Trials | Chr  | SNP                  | Position (bp) | Major/Minor allele | MAF | SNP effect | SE | P-value   |
|--------|------|----------------------|---------------|--------------------|-----|------------|----|-----------|
| col_IgG 1 | 1    | BovineHD0100036412   | 128805177     | A/G                | 0.281 | 11.960     | 2.964 | 4.20E-06  |
| col_IgG 1 | 1    | BTB-02067404         | 129170050     | A/G                | 0.253 | 12.451     | 3.067 | 6.02E-05  |
| col_IgG 23 | 23   | BovineHD2300015105   | 52008189      | A/G                | 0.431 | 11.301     | 2.818 | 4.38E-06  |
| col_IgG 23 | 23   | BovineHD2300015111   | 52023784      | A/G                | 0.430 | 11.358     | 2.813 | 2.15E-05  |
| col_IgG 24 | 24   | BovineHD2400012624   | 45892763      | C/A                | 0.467 | 11.042     | 2.752 | 2.70E-05  |
| col_IgG 24 | 24   | ARS-BFGL-NGS-18561   | 45933554      | G/A                | 0.452 | 11.211     | 2.728 | 5.45E-05  |
| col_IgG 24 | 24   | BovineHD2400012657   | 45972409      | A/G                | 0.366 | 13.095     | 2.851 | 3.97E-05  |
| col_IgG 24 | 24   | ARS-BFGL-NGS-20011   | 45973112      | G/A                | 0.364 | 13.160     | 2.860 | 5.40E-05  |
| col_IgG 24 | 24   | ARS-BFGL-NGS-106793  | 45998467      | A/C                | 0.378 | 11.689     | 2.785 | 4.92E-05  |
| col_IgG 24 | 24   | ARS-BFGL-NGS-44158   | 46027109      | G/A                | 0.377 | 11.869     | 2.794 | 6.06E-05  |
| col_IgG 24 | 24   | BovineHD2400012729   | 46241172      | A/G                | 0.351 | 11.496     | 2.845 | 5.32E-05  |
| col_IgG1 1  | 1    | BovineHD0100020052   | 70005456      | G/A                | 0.363 | 2.027      | 0.457 | 9.23E-06  |
| col_IgG1 3  | 3    | ARS-BFGL-NGS-105483  | 52329712      | A/G                | 0.457 | -1.770     | 0.437 | 5.07E-05  |
| col_IgG1 5  | 5    | BovineHD05000008574  | 29136394      | A/G                | 0.195 | -2.652     | 0.571 | 3.43E-06  |
| col_IgG1 6  | 6    | BovineHD0600010748   | 38816105      | G/A                | 0.312 | 1.900      | 0.472 | 5.72E-05  |
| col_IgG1 20 | 20   | BovineHD2000008079   | 27230044      | A/G                | 0.386 | -1.946     | 0.483 | 5.59E-05  |
| col_IgG1 22 | 22   | BovineHD2200013311   | 46250930      | G/A                | 0.407 | -2.011     | 0.493 | 4.50E-05  |
| col_IgG1 27 | 27   | ARS-BFGL-NGS-24322   | 19137853      | G/C                | 0.398 | -1.936     | 0.462 | 2.75E-05  |
| col_IgG2 | 3  | BTB-00163972 | 118145216 | A/G  | 0.324 | 0.147 | 0.035 | 2.88E-05 |
|----------|----|-------------|-----------|------|-------|-------|-------|----------|
| col_IgG2 | 20 | BovineHD2000021006 | 71896856 | G/A  | 0.273 | 0.223 | 0.037 | 1.13E-09 |
| col_IgG2 | 21 | BovineHD2100019547 | 66578213 | A/G  | 0.403 | -0.139 | 0.032 | 1.75E-05 |
| col_IgG2 | 21 | BovineHD2100020165 | 69062573 | G/A  | 0.399 | -0.135 | 0.032 | 2.72E-05 |
| col_IgG2 | 21 | BovineHD2100020225 | 69289258 | A/G  | 0.331 | 0.145  | 0.033 | 1.08E-05 |
| col_IgG2 | 21 | BovineHD2100020241 | 69357379 | A/C  | 0.427 | 0.139  | 0.033 | 2.87E-05 |
| col_IgG2 | 21 | ARS-BFGL-NGS-115062 | 69395154 | G/A  | 0.450 | 0.133  | 0.033 | 5.98E-05 |
| col_IgG2 | 21 | BovineHD2100020341 | 69673486 | A/G  | 0.361 | 0.148  | 0.032 | 5.34E-06 |
| col_IgG2 | 21 | BovineHD2100020413 | 69920970 | G/A  | 0.387 | 0.132  | 0.033 | 5.57E-05 |
| col_IgG2 | 21 | ARS-BFGL-NGS-2644 | 70608408 | A/G  | 0.350 | 0.155  | 0.034 | 5.00E-06 |
| col_IgG2 | 21 | BovineHD2100020676 | 70621565 | G/A  | 0.350 | 0.159  | 0.034 | 3.22E-06 |
| col_IgG2 | 21 | BovineHD2100020685 | 70655075 | A/G  | 0.327 | 0.174  | 0.035 | 5.78E-07 |
| col_IgG2 | 21 | BovineHD2100020687 | 70665452 | G/A  | 0.281 | -0.150 | 0.037 | 4.31E-05 |
| col_IgG2 | 21 | BovineHD2100020689 | 70672433 | A/G  | 0.379 | 0.142  | 0.034 | 3.35E-05 |
| col_IgG2 | 21 | BovineHD2100020696 | 70687439 | G/A  | 0.368 | 0.138  | 0.033 | 2.44E-05 |
| col_IgG2 | 21 | ARS-BFGL-NGS-73522 | 70702245 | G/A  | 0.347 | 0.152  | 0.034 | 7.68E-06 |
| col_IgG2 | 21 | BovineHD4100015389 | 70726088 | C/A  | 0.367 | -0.142 | 0.034 | 2.84E-05 |
| col_IgG2 | 21 | BovineHD2100020833 | 71318798 | A/G  | 0.286 | 0.169  | 0.035 | 1.11E-06 |
| col_IgG2 | 21 | BovineHD2100020837 | 71338809 | G/A  | 0.148 | 0.176  | 0.044 | 6.25E-05 |
| col_IgG2 | 21 | BovineHD2100020853 | 71389313 | G/A  | 0.248 | -0.167 | 0.039 | 1.55E-05 |
| col_IgA  | 1  | BovineHD0100012276 | 43138880 | A/C  | 0.425 | 0.129  | 0.031 | 3.61E-07 |
| col_IgA | 4   | BTB-00172973       | 31573521 | G/A | 0.267 | -0.140 | 0.033 | 2.87E-05 |
|--------|-----|-------------------|----------|-----|-------|--------|-------|----------|
| col_IgA | 5   | BovineHD0500021305| 74998613 | G/A | 0.334 | -0.138 | 0.031 | 9.66E-06 |
| col_IgM | 8   | BovineHD0800026113| 88021995 | G/A | 0.484 | 0.189  | 0.047 | 6.22E-05 |
| col_IgM | 10  | BovineHD1000019825| 69080826 | A/C | 0.162 | -0.305 | 0.060 | 4.14E-07 |
| col_IgM | 10  | BovineHD1000019983| 69718494 | C/A | 0.323 | -0.242 | 0.049 | 9.53E-07 |
| col_IgM | 10  | BovineHD1000020063| 70071886 | A/G | 0.429 | -0.212 | 0.048 | 1.01E-05 |
| col_IgM | 10  | Hapmap29285-BTA-162580| 69721690 | T/A | 0.356 | -0.201 | 0.049 | 3.68E-05 |
| col_IgM | 15  | ARS-BFGL-NGS-109495| 1492400 | A/T | 0.379 | -0.187 | 0.047 | 6.16E-05 |
| col_IgM | 17  | BTB-01973064       | 73349750 | A/G | 0.391 | 0.234  | 0.048 | 1.33E-06 |
| col_IgM | 17  | BovineHD1700021434| 73349278 | A/G | 0.457 | 0.227  | 0.047 | 1.35E-06 |
| col_IgM | 17  | BovineHD1700021122| 72466328 | G/A | 0.439 | -0.215 | 0.047 | 4.18E-06 |
| col_IgM | 17  | ARS-BFGL-NGS-5369 | 71841734 | A/C | 0.298 | -0.216 | 0.051 | 2.51E-05 |

\[ \text{a} \text{col}_{IgG}, \text{col}_{IgG1}, \text{col}_{IgG2}, \text{col}_{IgA} \text{ and } \text{col}_{IgM} \text{ represent concentration of } IgG, IgG1 \text{ IgG2, IgA and IgM in colostrum, respectively.} \]

\[ \text{b} \text{Cow chromosome number.} \]

\[ \text{c} \text{Minor allele frequency.} \]

\[ \text{d} \text{SNP effect.} \]

\[ \text{e} \text{standard error} \]
Table 3: The significant SNPs for concentrations of IgG, IgG1, IgG2, IgA and IgM in serum.

| Trials | Chr | SNP             | Position (bp) | Major/ minor allele | MAF | SNP effect | SE  | P-value   |
|--------|-----|-----------------|---------------|---------------------|-----|------------|-----|-----------|
| ser_IgG1 | 28 | BovineHD2800011781 | 41894086 | A/G                 | 0.362 | 0.083 | 0.018 | 6.88E-06  |
| ser_IgG1 | 19 | Hapmap57648-rs29022376 | 22709090 | G/A                 | 0.347 | -0.084 | 0.019 | 1.28E-05  |
| ser_IgG1 | 7  | BovineHD0700032048 | 109730181 | A/G                 | 0.445 | -0.075 | 0.018 | 2.89E-05  |
| ser_IgG1 | 19 | ARS-BFGL-NGS-20165 | 23561900 | C/A                 | 0.410 | -0.075 | 0.018 | 4.19E-05  |
| ser_IgG2 | 12 | BovineHD1200008568 | 28980452 | C/A                 | 0.323 | 0.251 | 0.055 | 5.84E-06  |
| ser_IgG2 | 12 | BovineHD1200008569 | 28987016 | A/G                 | 0.478 | -0.204 | 0.050 | 5.16E-05  |
| ser_IgG2 | 20 | BovineHD2000021006 | 71896856 | G/A                 | 0.274 | 0.413 | 0.057 | 5.04E-13  |
| ser_IgG2 | 21 | BovineHD2100018787 | 63926754 | A/G                 | 0.464 | -0.229 | 0.051 | 7.19E-06  |
| ser_IgG2 | 21 | BTA-24891-no-rs   | 63955841 | G/A                 | 0.357 | 0.214 | 0.052 | 4.18E-05  |
| ser_IgG2 | 21 | BovineHD2100019656 | 66910728 | G/A                 | 0.416 | 0.215 | 0.051 | 2.36E-05  |
| ser_IgG2 | 21 | ARS-BFGL-NGS-20339 | 67088847 | G/A                 | 0.333 | 0.240 | 0.052 | 4.71E-06  |
| ser_IgG2 | 21 | BovineHD2100019814 | 67542721 | A/C                 | 0.363 | 0.240 | 0.053 | 5.03E-06  |
| ser_IgG2 | 21 | BovineHD2100019834 | 67604077 | A/C                 | 0.147 | 0.289 | 0.071 | 5.19E-05  |
| ser_IgG2 | 21 | BovineHD4100015366 | 67786313 | A/G                 | 0.421 | 0.267 | 0.052 | 2.94E-07  |
| ser_IgG2 | 21 | BovineHD2100019888 | 67885290 | A/G                 | 0.217 | 0.246 | 0.060 | 4.21E-05  |
| ser_IgG2 | 21 | BovineHD2100019906 | 67946189 | A/G                 | 0.319 | 0.241 | 0.054 | 6.86E-06  |
| ser_IgG2 | 21 | ARS-BFGL-NGS-86477 | 68399787 | A/C                 | 0.482 | -0.213 | 0.052 | 4.06E-05  |
| ser_IgG2 | 21 | BovineHD2100020157 | 69009950 | A/C                 | 0.442 | 0.229 | 0.052 | 1.07E-05  |
| ser-lgG2  | 21 | BovineHD2100021033 | 69033145 | G/A | 0.201 | 0.331 | 0.062 | 7.14E-08 |
|----------|----|-------------------|---------|-----|-------|-------|-------|---------|
| ser-lgG2 | 21 | BovineHD2100020225 | 69289258 | A/G | 0.331 | 0.272 | 0.052 | 1.37E-07 |
| ser-lgG2 | 21 | BovineHD2100020232 | 69327116 | G/A | 0.349 | 0.271 | 0.051 | 1.32E-07 |
| ser-lgG2 | 21 | BovineHD2100020241 | 69357379 | A/C | 0.424 | 0.224 | 0.053 | 2.14E-05 |
| ser-lgG2 | 21 | ARS-BFGL-NGS-115062 | 69395154 | G/A | 0.450 | 0.231 | 0.053 | 1.17E-05 |
| ser-lgG2 | 21 | BovineHD2100020317 | 69613677 | G/A | 0.243 | 0.254 | 0.060 | 2.29E-05 |
| ser-lgG2 | 21 | BovineHD2100020341 | 69673486 | A/G | 0.366 | 0.299 | 0.051 | 4.21E-09 |
| ser-lgG2 | 21 | BovineHD2100020413 | 69920970 | G/A | 0.386 | 0.263 | 0.052 | 3.92E-07 |
| ser-lgG2 | 21 | ARS-BFGL-NGS-1345 | 69939350 | C/A | 0.420 | 0.232 | 0.052 | 7.32E-06 |
| ser-lgG2 | 21 | BovineHD2100020439 | 70000656 | A/G | 0.446 | 0.239 | 0.053 | 6.19E-06 |
| ser-lgG2 | 21 | ARS-USDA-AGIL-chr21-70182028-000470 | 70182028 | C/G | 0.376 | 0.212 | 0.052 | 5.34E-05 |
| ser-lgG2 | 21 | BovineHD2100020653 | 70537404 | A/G | 0.182 | 0.380 | 0.063 | 2.17E-09 |
| ser-lgG2 | 21 | BovineHD2100020670 | 70592463 | A/C | 0.415 | 0.251 | 0.052 | 1.52E-06 |
| ser-lgG2 | 21 | ARS-BFGL-NGS-2644 | 70608408 | A/G | 0.353 | 0.330 | 0.053 | 6.66E-10 |
| ser-lgG2 | 21 | BovineHD2100020676 | 70621565 | G/A | 0.353 | 0.343 | 0.054 | 1.74E-10 |
| ser-lgG2 | 21 | BovineHD2100020685 | 70655075 | A/G | 0.329 | 0.342 | 0.055 | 4.05E-10 |
| ser-lgG2 | 21 | BovineHD2100020689 | 70672433 | A/G | 0.381 | 0.256 | 0.054 | 1.73E-06 |
| ser-lgG2 | 21 | BovineHD2100020696 | 70687439 | G/A | 0.370 | 0.296 | 0.051 | 8.63E-09 |
| ser-lgG2 | 21 | ARS-BFGL-NGS-73522 | 70702245 | G/A | 0.349 | 0.287 | 0.053 | 6.29E-08 |
| ser-lgG2 | 21 | BovineHD2100020833 | 71318798 | A/G | 0.286 | 0.322 | 0.055 | 6.15E-09 |
| ser-lgG2 | 21 | BovineHD2100020883 | 71479429 | A/G | 0.165 | 0.400 | 0.068 | 3.18E-09 |
| ser_IgA | 2 | ARS-BFGL-NGS-100214 | 132821698 | A/G | 0.139 | -0.142 | 0.033 | 1.76E-05 |
|---------|---|----------------------|-----------|-----|-------|--------|-------|----------|
| ser_IgA | 15 | BTA-91367-no-rs | 60316301 | G/A | 0.464 | -0.099 | 0.024 | 2.52E-05 |
| ser_IgA | 15 | BovineHD1500016066 | 55611146 | G/A | 0.477 | -0.101 | 0.025 | 4.69E-05 |
| ser_IgA | 15 | BovineHD1500016061 | 55590128 | G/A | 0.476 | -0.100 | 0.025 | 5.43E-05 |
| ser_IgA | 18 | BovineHD1800015730 | 53613268 | G/A | 0.133 | -0.150 | 0.037 | 5.23E-05 |
| ser_IgA | 25 | BovineHD2500007939 | 28524826 | A/G | 0.387 | -0.106 | 0.025 | 1.88E-05 |
| ser_IgA | 25 | BovineHD2500007948 | 28580391 | A/G | 0.388 | -0.107 | 0.025 | 1.96E-05 |
| ser_IgA | 29 | BovineHD2900006144 | 21381114 | A/G | 0.201 | -0.115 | 0.028 | 5.55E-05 |
| ser_IgM | 4 | BovineHD0400027460 | 98194248 | A/G | 0.115 | 0.125 | 0.031 | 6.29E-05 |
| ser_IgM | 6 | BovineHD0600009523 | 34015077 | G/A | 0.270 | -0.101 | 0.022 | 3.73E-06 |
| ser_IgM | 7 | ARS-BFGL-NGS-12159 | 19220954 | A/C | 0.433 | -0.084 | 0.020 | 1.56E-05 |
| ser_IgM | 7 | BovineHD0700005341 | 19183296 | A/G | 0.432 | -0.081 | 0.019 | 3.45E-05 |
| ser_IgM | 17 | BovineHD1700021706 | 74223510 | G/A | 0.170 | -0.123 | 0.026 | 2.74E-06 |

\(^a\)ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represent concentration of IgG1 IgG2, IgA and IgM in serum, respectively.

\(^b\)Cow Chromosome number.

\(^c\)Minor allele frequency.

\(^d\)SNP effect.

\(^e\)standard error
Table 4: The list of candidate genes nearby the SNPs associated with IgG, IgG1, IgG2, IgA and IgM in the colostrum and serum.

| Gene ID                  | Gene Name | Chr | Gene Start   | Gene End   | Traits          |
|--------------------------|-----------|-----|--------------|------------|-----------------|
| ENSBTAG000000008424      | ABR       | 19  | 22287405     | 22431563   | ser_IgG1        |
| ENSBTAG000000008423      | TIMM22    | 19  | 22432982     | 22439070   | ser_IgG1        |
| ENSBTAG00000005665       | CRK       | 19  | 23139102     | 23159120   | ser_IgG1        |
| ENSBTAG0000001332        | MYO1C     | 19  | 23168435     | 23191280   | ser_IgG1        |
| ENSBTAG00000038889       | RILP      | 19  | 23320981     | 23323996   | ser_IgG1        |
| ENSBTAG00000020859       | SERPINF2  | 19  | 23403657     | 23411269   | ser_IgG1        |
| ENSBTAG00000017636       | AKT1      | 21  | 70878138     | 70895537   | col_IgG2, ser_IgG2 |
| ENSBTAG00000018019       | BCL11B    | 21  | 65841997     | 65938304   | col_IgG2, ser_IgG2 |
| ENSBTAG00000026913       | HHIPL1    | 21  | 66284693     | 66314293   | col_IgG2, ser_IgG2 |
| ENSBTAG00000016598       | DYNC1H1   | 21  | 68507134     | 68568024   | col_IgG2, ser_IgG2 |
| ENSBTAG00000006270       | HSP90AA1  | 21  | 68595873     | 68601204   | col_IgG2, ser_IgG2 |
| ENSBTAG00000013475       | TRAF3     | 21  | 69228047     | 69254153   | col_IgG2, ser_IgG2 |
| ENSBTAG00000017299       | KLC1      | 21  | 69896140     | 69957173   | col_IgG2, ser_IgG2 |
| ENSBTAG00000014921       | IL6       | 4   | 31578311     | 31582667   | col_IgA         |
| ENSBTAG00000020535       | PYCARD    | 25  | 27541409     | 27542740   | ser_IgA         |
| ENSBTAG00000047238       | ITGAM     | 25  | 27602222     | 27639821   | ser_IgA         |
| ENSBTAG0000002391        | TGFBI1    | 25  | 27760811     | 27766954   | ser_IgA         |
| ENSBTAG0000000704        | GUSB      | 25  | 28162569     | 28174955   | ser_IgA         |
| ENSBTAG00000037489       | CRCP      | 25  | 28255764     | 28294982   | ser_IgA         |
| ENSBTAG0000003842        | RABGEF1   | 25  | 28509094     | 28534880   | ser_IgA         |
| ENSBTAG00000004051       | SBDS      | 25  | 28630825     | 28636300   | ser_IgA         |

aCow chromosome number.

bGene contain SNP significantly associated with concentration of IgG, IgG1 IgG2, IgA and IgM in colostrum and serum, respectively. col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represent concentration of IgG, IgG1 IgG2, IgA and IgM in colostrum, ser_IgG, ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represent concentration of IgG, IgG1 IgG2, IgA and IgM in serum, respectively.

Figures
Figure 1

Q-Q plots of the observed P-values for concentrations of IgG, IgG1, IgG2, IgA and IgM in the colostrum and serum. The Q-Q plots show the observed $-\log_{10}$-transformed P-values (y-axis) and the expected $-\log_{10}$-transformed P-values (x-axis).

Figure 2

Manhattan plots of the observed P-values for concentrations of IgG, IgG1, IgG2, IgA and IgM in the colostrum and serum. The Manhattan plots indicate $-\log_{10}$ (P-values) for genome-wide SNPs (y-axis) plotted against their respective positions on each chromosome (x-axis), the horizontal red and red dashed lines in the Manhattan plots indicate the genome-wide ($3.16E^{-6}$) and suggestive significance ($6.32E^{-5}$) thresholds respectively.
Figure 3

Haplotype blocks formed by the significant SNPs. (A) Indicate a haplotype block composed of the significant SNPs for IgG concentration in colostrum. (B) Indicate a haplotype block composed of the significant SNPs for IgG2 concentration in colostrum. (C) Indicate a haplotype block composed of the significant SNPs for IgG2 concentration in serum. The black lines mark the identified blocks.

Supplementary Files

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

- TableS1.xlsx
- TableS2.xlsx