Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genome-Wide Association Studies for Immunoglobulin Concentrations 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 Univerdity
Bo Han
China Agricultural University
Yaofeng Zhao
China Agricultural University
Shengli Zhang
China Agricultural University
Dongxiao Sun (sundx@cau.edu.cn)
China Agricultural University https://orcid.org/0000-0002-7512-5946

Research article

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

Posted Date: September 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-42791/v1

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

Version of Record: A version of this preprint was published at BMC Genomics on January 10th, 2022. See the published version at https://doi.org/10.1186/s12864-021-08250-5.

Abstract

Background

The early death and health problems of calves caused substantial economic losses in dairy industry. As the immune system has not been fully developed in the neonates, transport of passive immune substance such as immunoglobulins (Ig) from mothers to newborn calves is essential in protecting neonates from infections in their early life. Therefore, concentrations of immunoglobulins in the colostrum and serum of dairy cows are critical traits when estimating potential disease resistance of its offspring.

Results

Colostrum, blood and hair follicle samples were collected from the 620 Chinese Holstein cows within 24 hours after calving. The concentration of total IgG, IgG1, IgG2, IgA and IgM in both colostrum and serum were detected via ELISA methods. With GCTA software, genome-wide association studies (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. As a result, 20, 1 and 4 significant SNPs were detected associated with the concentrations of IgG2, IgA and IgM at genome-wide level ($P < 3.16E{-}6$); 11, 11, 35, 11 and 10 significant SNPs were identified associated with total IgG, IgG1, IgG2, IgA and IgM at suggestive level ($P < 6.32E{-}5$). Such SNPs were located in or proximate to (±1 Mb) 1,083 genes, which were functionally implicated in biological processes and pathways, such as immune response, negative regulation of immunoglobulin secretion, Fc-epsilon receptor and NF-kappaB signaling pathways. By combining the biological functions and the known QTL data for immune traits in bovine, 21 promising candidate functional genes were identified for immunoglobulins concentrations in colostrum and serum in dairy cattle, they were *ABR, TIMM22, CRK, MYO1C, RILP, SERPINF2, AKT1, BCL11B, HHIPL1, DYNC1H1, HSP90AA1, TRAF3, KLC1, IL6, PYCARD, ITGAM, TGFB1I1, GUSB, CRCP, RABGEF1* and *SBDS*.

Conclusions

In this study, we identified 21 candidate genes related to concentrations of immunoglobulins in colostrum and serum in dairy cattle by performing GWASs. Our findings provide a groundwork for unraveling the key genes and causal mutations affecting immunoglobulins levels in colostrum and important information for genetic improvement of such traits in dairy cattle.

Background

The early survival rate and health of calves is one of the important factors affecting the production efficiency of dairy industry. The sufficient intake of high-quality colostrum determines physical conditions, growth rate and early survival of newborn dairy calves [1, 2]. As the immune system has not been fully developed in the neonates, transport of passive immune substance such as immunoglobulins (Ig) from mothers to newborn calves is essential in protecting neonates from infections in their early life. Immunoglobulins are the major protein components of colostrum, comprising 70–80% of the total protein content, while in mature milk, immunoglobulins constitute only 1–2% [3, 4]. There are three major immunoglobulins in bovine serum and milk: IgG, IgM and IgA, with IgG consisting of two subclasses (IgG1 and IgG2). IgG1 accounts for over 75%

(46.4 mg/ml), and IgM (6.8 mg/ml), IgA (5.4 mg/ml) and IgG2 (2.9 mg/ml) were followed successively [3]. Immunoglobulins are produced by B1-cells 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 [5]. Hence, delineation of the genetic architecture underlying the concentrations of immunoglobulins in cows' colostrum is important for identifying ways to improve survival rate of neonatal calves in dairy cattle.

Concentrations of immunoglobulins are typical quantitative characteristics controlled by multiple QTLs and polygenes. Heritability estimates for IgG concentrations in blood ranged from 0.27 to 0.64 in human [6–9]. Similarly, the heritability estimates for immunoglobulins in serum and milk 0.08–0.45 in dairy cattle [10–13], in which heritability of IgM (0.18–0.45) was broadly high, while estimates for IgG was relatively lower (0.08–0.31); compared with concentrations of immunoglobulins in milk (0.08–0.23), the heritability estimated in serum was generally higher (0.15–0.25) [10].

Genome-wide association studies (GWASs) have been performed for immunoglobulins in serum or mature milk. The first GWAS based on 2,247 individuals from four European cohorts (CROATIA-Vis, CROATIA-Korcula, Orkney Complex Disease Study and Northern Swedish Population Health Study) identified 9 genome-wide significant loci associate with IgG glycosylation and 4 out of them contained genes encoding glycosyltransferases [14]. Another GWAS for IgG 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 [15]. In pigs, 2 genome-wide and 4 chromosome-wide significant SNPs were detected for IgG blocking percentage to CSF virus in serum by performing GWAS [16]. Especially, a GWAS for blood natural antibodies in Canadian Holstein cows identified 23 SNPs that were significantly associated with IgG concentration at genome-wide level [17]. Another GWAS for milk natural antibodies in Dutch Holstein-Friesian cattle identified some significant SNPs for IgG1 and IgM with candidate genes on Bos taurus autosome (BTA) 3, 17, 18, and 21 that related to immunoglobulin structure and early B cell development [18]. However there is few study on gene identification for concentrations of immunoglobulins in colostrum in dairy cattle so far. In addition, in bovine colostrum, the immunoglobulins were found mainly derived from serum [19, 20]. Hence, investigation on the concentrations of immunoglobulins in both colostrum and serum can better undestand the genetic architecture underlying colostrum immunoglobulin traits. Here, we conducted genome wide association studies for the concentrations of immunoglobulin components in colostrum and strum in a Chinese Holstein population in order to identify the functional genes contributed to the phenotypic variation of colostrum immunoglobulins and provide molecular information for genetically improving such traits to increase calves' disease-resistance.

Methods

Animals and phenotypes

The animals used in this study consists of 620 Chinese Holstein cows who were the daughters of 44 sires with average 13.78 cows per sire and were collected from 10 dairy farms in the Beijing Dairy Cattle Center and the Beijing Sunlon Livestock Development Company Limited. All the cows were on average parity of 2.52,

ranging from 1 to 4. The blood serum and colostrum samples were taken from each cow during the first milking within 24 h after calving for measurement of immunoglobulins. Hair follicle samples were collected from each animal for SNP chip genotyping as well. The whole procedure for collection of the samples (blood, hair and colostrum) was implemented in strict accordance with the protocol approved by the Animal Welfare Committee of China Agricultural University (Permit number: DK996). The animals used in this study were all released to their own population for normal production after sample collection.

We used the commercially available, ELISA kits to measure the levels of immunoglobulins of each colostrum sample and serum sample, including total 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). For further statistical analysis, the phenotypic values for the concentrations of total IgG, IgG1 and IgG2 in colostrum or serum were square root transformed to fit a normal distribution, simultaneously phenotypes for IgA and IgM concentrations were log-transformed corrected.

Genotyping and quality control

Genomic DNA was extracted from the hair follicle samples with the QIAamp® DNA Mini Kit (QIAGEN, Valencia, CA, USA) for genotyping. A total of 588 individuals were genotyped with the GeneSeek GGP_HDv3 chip (including 140,668 SNP markers: GeneSeek, Lincoln, NE, USA) and 32 cows were genotyped with the Illumina 50K chip (including 54,609 SNP markers: Illumina, San Diego, CA, USA). To make full use of the SNP information originating from the GeneSeek 150K chip, we performed genotype imputation for the 32 cows genotyped with 50 k chip using BEAGLE 3.0.4 software with default option[21]. Allelic R^2 was used to evaluate the imputation accuracy based on the genotype probabilities. Then, we implemented PLINK 1.90 for quality control and removed samples with call rate of genotyping < 90%, SNPs with call rates < 90%, minor allele frequencies (MAF) < 0.1 and Hardy-Weinberg equilibrium (HWE) P values < 10^{-6} [22, 23]. The SNPs positions were mapped using the Bos taurus UMD 3.1 assembly. Finally, after quality control, 620 individuals with 88,934 SNPs remained for the GWASs (Additional files 1 and 2).

Statistical analysis

Mixed model based single locus regression analyses (MMRA)

We performed single-SNP association analysis for the individual phenotype in GCTA 1.90.2 with the following mixed linear model:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{g} + \mathbf{A}\sigma_g^2 + \mathbf{I}\sigma_\varepsilon^2$$

Where $\mathbf{y}$ 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; $\boldsymbol{\beta}$ is the vector of fixed effects, including herd, parity and season of calving, $\mathbf{X}$ is an incidence matrix relating elements of $\boldsymbol{\beta}$ to $\mathbf{y}$; $\mathbf{g}$ is an n × 1 vector of the total genetic effects of the individuals with $\mathbf{g} \sim N(0, \mathbf{A}\sigma_g^2)$, and $\mathbf{A}$ is interpreted as the genetic relationship matrix (GRM) between individuals. We can therefore estimate σ_g^2 by the restricted maximum likelihood (REML) approach,

relying on the GRM estimated from all the SNPs [24]. **I** is an n × 1 identity matrix, and ε is a vector of residual effects with $\varepsilon \sim N(0, I\sigma_{\varepsilon}^2)$.

The existence of linkage disequilibrium (LD) of SNPs in every chromosome may lead to over-correction when using Bonferroni adjustments [25, 26]. Hence, we used effectively independent test number to define the thresholds for genome-wide significant/suggestive associations based on the assessed number of independent markers and linkage disequilibrium blocks for markers on every chromosome [27]. We calculated the LD extend between two adjacent SNPs with PLINK 1.90 and the pairwise r^2 values greater than 0.40 were defined as a linkage disequilibrium block [28].

Population stratification can result in spurious association findings in a GWAS [29]. Thus, we calculated genomic inflation factor (λ) and also depicted quantile-quantile (Q-Q) plot to assess stratification in our study population using GENABEL and qqman packages in R 3.6.0. Further, we performed LD analysis for the identified significantly SNPs using Haploview v4.2 [30]. The block was defined according to the solid spin algorithm by the criteria of Gabriel [31].

Identification of candidate genes

For the identified significant SNPs associated with the concentrations of immunoglobulins, we searched for the functional genes that were located less than 1 Mb to the significant SNPs or contained the significant SNPs based on the UMD 3.1 assembly with BioMart tool (http://asia.ensembl.org/biomart/martview) in Ensemble database (Ensemble Genes Release 92). To know about the biological functions of these genes, we implemented Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis with KOBAS 3.0 software [32]. In addition, we also compared the physical position of these functional genes with the reported quantitative traits loci (QTLs) for immune capacity traits in the Cattle QTL database (https://www.animalgenome.org/cgi-bin/QTLdb/BT/index).

Results

Statistics of phenotypes

In this study, we measured the concentrations of immunoglobulins in both colostrum and serum for 620 Chinese Holstein cows. As a result, a total of 10 traits were acquired, including concentrations of total IgG, IgG1, IgG2, IgA and IgM. Means and the corresponding standard deviations for the original and corrected phenotypic values were shown in Table 1.

Table 1
Means and standard deviations for the original and corrected concentrations of immunoglobulins in colostrum and serum.

Traits	N	Mean	SD	Min	Max	Transformation	Mean	SD	Min	Max
col_IgG	540	33.70	43.17	1.16	342.05	col_IgG_sqrt	5.19	2.61	1.08	18.49
col_IgG1	542	15.09	7.78	0.24	44.83	col_IgG1_sqrt	3.74	1.05	0.49	6.70
col_IgG2	606	3.2	1.94	0.07	11.04	col_IgG2_sqrt	1.7	0.54	0.26	3.32
ser_IgG	607	8.34	3.37	0.68	22.9	ser_IgG_sqrt	2.83	0.58	0.82	4.79
ser_IgG1	612	1.04	0.74	0.01	7.65	ser_IgG1_sqrt	1.04	0.31	0.09	2.77
ser_IgG2	614	13.79	6.51	0.69	42.36	ser_IgG2_sqrt	3.61	0.86	0.83	6.51
col_IgA	613	3.11	3.59	0.01	51.18	col_IgA_log	0.28	0.49	-1.91	1.71
col_IgM	613	5.47	3.7	0.09	36.41	col_IgM_log	0.63	0.35	-1.06	1.56
ser_IgA	613	0.23	0.17	0.01	1.21	ser_IgA_log	-0.73	0.39	-5.4	0.08
ser_IgM	613	2.25	1.68	0.01	14.93	ser_IgM_log	0.24	0.34	-2.34	1.17

N, sample number; Mean, arithmetic mean; SD, standard deviation; Min, minimum; Max, maximum; col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represented the concentration of total IgG, IgG1, IgG2, IgA and IgM in colostrum, respectively; ser_IgG, ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represented the concentration of total IgG, IgG1, IgG2, IgA and IgM in serum, respectively; _sqrt, corrected phenotypes by square root transformation; _log, corrected phenotypes by log transformation.

Genotype imputation

After genotype imputation with BEAGLE 3.0.4, a total of 109,009 SNPs with $R^2 > 0.3$ in the 32 imputed cows was retained, in which the average allelic R^2 of imputed genotypes was 74.8% [33]. Out of these, 88,934 SNPs were included in the further association analysis after quality control.

Genome-wide association study

After LD analysis, a total of 15,820 effectively independent tests number were suggested. Thus, the threshold P-value for genome-wide significant association was set for 3.16E–6 (0.05/15,820) and that for suggestive significant association was 6.32E–5 (1/15,820) [27]. Based on the QQ plots (Figs. 1 and 2) and the estimated inflation factor (λ) of 0.98–1.02 for all traits, no population stratification was observed.

With GCTA 1.90.2, we performed the genome-wide association studies for the 10 traits. For the concentrations of immunoglobulins in colostrum (Fig. 1 and Table 2), 3 genome-wide significant SNPs for IgG2 were detected (P < 3.16E-6), including 1 SNP on BTA20 and 2 SNPs in BTA21; at suggestive level, a total of 11, 7 and 17 SNPs (P < 6.32E-5) were detected for total IgG, IgG1 and IgG2, respectively, locating on BTA1 (3 SNPs), 3 (2 SNPs), 5 (1 SNP), 6 (1 SNP), 20 (2 SNPs), 21 (15 SNPs), 23 (2 SNPs), 24 (7 SNPs) and 27 (2 SNPs). In addition, 2 SNPs in BTA10 and 2 SNPs in BTA17 associated with concentration of IgM were detected at genome-wide significant level (P < 3.16E-6); 3 and 6 significant SNPs (P < 6.32E-5) were detected for IgA and

IgM at suggestive level, respectively, that distributed in BTA1 (1 SNP), 4 (1 SNP), 5 (1 SNP), 8 (1 SNP), 10 (2 SNPs), 15 (1 SNP) and 17 (2 SNPs).

For the concentrations of immunoglobulins in serum (Fig. 2 and Table 3), at genome-wide significant level, 16 SNPs in BTA21 and 1 SNP in BTA20 were identified associated with IgG2 ($P < 3.16E-6$); at suggestive significant level, 4 SNPs and 18 SNPs were detected for IgG1 and IgG2 ($P < 6.32E-5$), locating on BTA7 (1 SNP), 12 (2 SNPs), 19 (2 SNPs), 20 (1 SNP) 21 (15 SNPs) and 28 (1 SNP); at both level, total IgG did not show any significant associations. Simultaneously, for IgA and IgM in serum, 1 SNP in BTA17 was significantly associated with IgA at genome-wide level ($P < 3.16E-6$); 8 and 4 SNPs were identified for IgA and IgM at suggestive significant level ($P < 6.32E-5$), that spread in BTA2 (1 SNP), 15 (3 SNPs), 18 (1 SNP), 25 (2 SNP), 29 (1 SNP), 4 (1 SNP), 6 (1 SNP) and 7 (2 SNPs).

Through LD analysis, we observed that the 7 significant SNPs for total IgG concentration in colostrum in BTA24 (45.89 ~ 46.24 Mb) generated a haplotype block ($D'=86 \sim 100$, Fig. 3A); for IgG2 in colostrum, 18 significant SNPs in BTA21 (66.58 ~ 71.39 Mb) formed 3 haplotype blocks: block 1 (14 kb, 2 SNPs), block 2 (17 kb, 3 SNPs) and block 3 (14 kb, 2 SNPs) ($D'=97 \sim 100$, Fig. 3B); for IgG2 in serum, 3 haplotype blocks were found form 32 significant SNPs on BAT 21 (64.93 ~ 72.48 Mb) that were covered by blocks of IgG2 in colostrum, including block 1 (18 kb, 2 SNPs), block 2 (62 kb, 4 SNPs) and block 3 (14 kb, 2 SNPs) ($D'=96 \sim 100$, Fig. 3C).

Table 2
The significant SNPs for concentrations of total IgG, IgG1, IgG2, IgA and IgM in colostrum.

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	P-value
col_IgG	1	BovineHD0100036412	128805177	A/G	0.281	11.960	2.964	4.20E-06
col_IgG	1	BTB-02067404	129170050	A/G	0.253	12.451	3.067	6.02E-05
col_IgG	23	BovineHD2300015105	52008189	A/G	0.431	11.301	2.818	4.38E-06
col_IgG	23	BovineHD2300015111	52023784	A/G	0.430	11.358	2.813	2.15E-05
col_IgG	24	BovineHD2400012624	45892763	C/A	0.467	11.042	2.752	2.70E-05
col_IgG	24	ARS-BFGL-NGS-18561	45933554	G/A	0.452	11.211	2.728	5.45E-05
col_IgG	24	BovineHD2400012657	45972409	A/G	0.366	13.095	2.851	3.97E-05
col_IgG	24	ARS-BFGL-NGS-20011	45973112	G/A	0.364	13.160	2.860	5.40E-05
col_IgG	24	ARS-BFGL-NGS-106793	45998467	A/C	0.378	11.689	2.785	4.92E-05
col_IgG	24	ARS-BFGL-NGS-44158	46027109	G/A	0.377	11.869	2.794	6.06E-05
col_IgG	24	BovineHD2400012729	46241172	A/G	0.351	11.496	2.845	5.32E-05
col_IgG1	1	BovineHD0100020052	70005456	G/A	0.363	2.027	0.457	9.23E-06
col_IgG1	3	ARS-BFGL-NGS-105483	52329712	A/G	0.457	-1.770	0.437	5.07E-05
col_IgG1	5	BovineHD0500008574	29136394	A/G	0.195	-2.652	0.571	3.43E-06

[a]col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represented concentration of total IgG, IgG1 IgG2, IgA and IgM in colostrum, respectively.

[b]Cow chromosome number.

[c]Minor allele frequency.

[d]standard error

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	P-value
col_IgG1	6	BovineHD0600010748	38816105	G/A	0.312	1.900	0.472	5.72E-05
col_IgG1	20	BovineHD2000008079	27230044	A/G	0.386	-1.946	0.483	5.59E-05
col_IgG1	22	BovineHD2200013311	46250930	G/A	0.407	-2.011	0.493	4.50E-05
col_IgG1	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

[a]col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represented concentration of total IgG, IgG1 IgG2, IgA and IgM in colostrum, respectively.

[b]Cow chromosome number.

[c]Minor allele frequency.

[d]standard error

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	P-value
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-05
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

[a]col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represented concentration of total IgG, IgG1 IgG2, IgA and IgM in colostrum, respectively.

[b]Cow chromosome number.

[c]Minor allele frequency.

[d]standard error

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	P-value
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

[a]col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represented concentration of total IgG, IgG1 IgG2, IgA and IgM in colostrum, respectively.

[b]Cow chromosome number.

[c]Minor allele frequency.

[d]standard error

Table 3
The significant SNPs for concentrations of total IgG, IgG1, IgG2, IgA and IgM in serum.

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	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

[a]ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represented concentration of IgG1 IgG2, IgA and IgM in serum, respectively.

[b]Cow Chromosome number.

[c]Minor allele frequency.

[d]standard error

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	P-value
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_IgG2	21	BovineHD2100021033	69033145	G/A	0.201	0.331	0.062	7.14E-08
ser_IgG2	21	BovineHD2100020225	69289258	A/G	0.331	0.272	0.052	1.37E-07
ser_IgG2	21	BovineHD2100020232	69327116	G/A	0.349	0.271	0.051	1.32E-07
ser_IgG2	21	BovineHD2100020241	69357379	A/C	0.424	0.224	0.053	2.14E-05
ser_IgG2	21	ARS-BFGL-NGS-115062	69395154	G/A	0.450	0.231	0.053	1.17E-05
ser_IgG2	21	BovineHD2100020317	69613677	G/A	0.243	0.254	0.060	2.29E-05
ser_IgG2	21	BovineHD2100020341	69673486	A/G	0.366	0.299	0.051	4.21E-09
ser_IgG2	21	BovineHD2100020413	69920970	G/A	0.386	0.263	0.052	3.92E-07
ser_IgG2	21	ARS-BFGL-NGS-1345	69939350	C/A	0.420	0.232	0.052	7.32E-06
ser_IgG2	21	BovineHD2100020439	70000656	A/G	0.446	0.239	0.053	6.19E-06
ser_IgG2	21	ARS-USDA-AGIL-chr21-70182028-000470	70182028	C/G	0.376	0.212	0.052	5.34E-05

[a]ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represented concentration of IgG1 IgG2, IgA and IgM in serum, respectively.

[b]Cow Chromosome number.

[c]Minor allele frequency.

[d]standard error

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	P-value
ser_IgG2	21	BovineHD2100020653	70537404	A/G	0.182	0.380	0.063	2.17E-09
ser_IgG2	21	BovineHD2100020670	70592463	A/C	0.415	0.251	0.052	1.52E-06
ser_IgG2	21	ARS-BFGL-NGS-2644	70608408	A/G	0.353	0.330	0.053	6.66E-10
ser_IgG2	21	BovineHD2100020676	70621565	G/A	0.353	0.343	0.054	1.74E-10
ser_IgG2	21	BovineHD2100020685	70655075	A/G	0.329	0.342	0.055	4.05E-10
ser_IgG2	21	BovineHD2100020689	70672433	A/G	0.381	0.256	0.054	1.73E-06
ser_IgG2	21	BovineHD2100020696	70687439	G/A	0.370	0.296	0.051	8.63E-09
ser_IgG2	21	ARS-BFGL-NGS-73522	70702245	G/A	0.349	0.287	0.053	6.29E-08
ser_IgG2	21	BovineHD2100020833	71318798	A/G	0.286	0.322	0.055	6.15E-09
ser_IgG2	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

[a]ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represented concentration of IgG1 IgG2, IgA and IgM in serum, respectively.

[b]Cow Chromosome number.

[c]Minor allele frequency.

[d]standard error

Traits[a]	Chr[b]	SNP name	Position (bp)	Major/ minor allele	MAF[c]	SNP effect	SE[d]	P-value
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 represented concentration of IgG1 IgG2, IgA and IgM in serum, respectively.

[b]Cow Chromosome number.

[c]Minor allele frequency.

[d]standard error

Candidate genes and function analysis

Appling BioMart in the Ensembl database (Ensembl Genes 92), we obtained 49, 221, 437, 252 and 408 functional genes that contained or were near to (within 1 Mb) the significant SNPs associated with total IgG, IgG1, IgG2, IgA and IgM traits, respectively. After removing the duplicates of genes, a total of 1,083 candidate genes was remained, including 706 protein-coding genes, 60 miRNA genes, 45 spliceosomal RNAs, 76 small nucleolar RNAs and 196 novel genes (Additional file 3: Table S1).

To further know about the biological functions of these candidate genes, we performed GO and KEGG analysis and observed that 151 genes were enriched in immune related biological processes and 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). Simultaneously, we compared the physical positions of the 1,083 genes with the peak of the known QTLs that have been shown associated with immune capacity in dairy cattle (Cattle QTLdb), including

IgG level, FMDV peptide-induced cell proliferation and ConA-induced cell proliferation. Consequently, 181 genes were found located within the QTL regions with distance to the peak positions of less than 1.0 cM.

Integrating the results of GO/KEGG and QTL data, 21 overlapping genes were considered as promising candidates for the concentrations of immunoglobulins in colostrum and serum (Table 4). They were ABR 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*), HHIP 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*).

Table 4
The list of candidate genes contained or nearby the significant SNPs associated with total IgG, IgG1, IgG2, IgA and IgM in the colostrum and serum.

Gene ID	Gene Name	Chr[a]	Gene Start[b]	Gene End[b]	Traits
ENSBTAG00000008424	*ABR*	19	22287405	22431563	ser_IgG1
ENSBTAG00000008423	*TIMM22*	19	22432982	22439070	ser_IgG1
ENSBTAG00000005665	*CRK*	19	23139102	23159120	ser_IgG1
ENSBTAG00000001332	*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
ENSBTAG00000002391	*TGFB1I1*	25	27760811	27766954	ser_IgA
ENSBTAG00000000704	*GUSB*	25	28162569	28174955	ser_IgA
ENSBTAG00000037489	*CRCP*	25	28255764	28294982	ser_IgA
ENSBTAG00000003842	*RABGEF1*	25	28509094	28534880	ser_IgA
ENSBTAG00000004051	*SBDS*	25	28630825	28636300	ser_IgA

[a]Cow chromosome number.

[b]The position of gene was based on the UMD 3.1 assembly.

col_IgG, col_IgG1, col_IgG2, col_IgA and col_IgM represented concentration of total IgG, IgG1 IgG2, IgA and IgM in colostrum; ser_IgG, ser_IgG1, ser_IgG2, ser_IgA and ser_IgM represented concentration of total IgG, IgG1 IgG2, IgA and IgM in serum.

Discussion

In this study, we identified the chromosome regions influencing the concentrations of immunoglobulins in colostrum and serum in dairy cattle by performing GWASs with high density SNP genotypes. Consequently, we detected 11, 11, 55, 77, 11 and 15 SNPs associated with the total IgG, IgG1, IgA and IgM at genome-wide and suggestive levels, respectively. To our knowledge, this is the first investigation on the genetic architecture for colostrum immunoglobulins in dairy cattle.

In general, a genomic inflation factor λ of < 1.05 demonstrates no population stratification [34]. In this study, the calculated λ values were 0.98–1.02 for concentrations of IgG, IgM and IgA traits in colostrum and serum suggesting population stratification was well controlled. In present study, the significant SNPs for the total IgG and IgG1 concentrations in colostrum and strum distributed on multiple chromosomes including BTA1, 3, 5, 6, 7, 19, 20, 22, 23 24, 27 and 28, while two previous GWASs in Canadian and Dutch Holstein populations observed that the significant SNPs for IgG in serum and IgG1 in mature milk were mainly located in BTA21 (55.5 to 70.6 Mb and 66.0 to 71.6 Mb) [17, 18]. Such inconsistences were most likely due to the different sampling time. We collected the colostrum and serum samples within 24 hours after calving, whereas their serum sample collected during 0 to 498 days in lactation [17] and the mature milk during 66 to 263 days in lactation [18] were used. For IgM concentrations in colostrum and serum, we detected 5 significant SNPs on BTA17 (71.8–74.2 Mb), overlapping 2 significant SNPs on BTA17 (72.5 to 73.6 Mb) identified for IgM in mature milk in a previous GWAS in Dutch Holstein populations [17]. For IgA, no related SNPs have been described in previous studies [17, 18].

Candidate genes

Combing the biological functions of the 1,083 functional genes that contained or were closed to the significant SNPs with less than 1 Mb and the known QTL data for immune traits in bovine, 21 candidate genes were selected for the immunoglobulins concentrations in colostrum and serum in dairy cattle.

For IgG1, 6 candidate genes within 22.3 ~ 23.4 Mb on BTA19 were identified. The protein encoded by the *ABR* gene is one of the Rho family members which is key regulator of phagocytic leukocyte function [35]. ABR curbed specific functions of mature tissue innate immune cells and regulated the inflammatory processes [36]. The protein encoded by *TIMM22* is a subunit of Tim22. Granzyme B depended on Tim22 to enter the mitochondria to induce apoptosis [37]. *CRK* encodes a member of an adapter protein family which was required for effector T cell trafficking into sites of inflammation [38]. In a murine model, CRK family proteins selectively regulate T cell adhesion and migration at effector sites and can act as the potential therapeutic targets for preventing graft versus-host disease (GVHD) [38]. *MYO1C* encodes a member of the unconventional myosin protein family that plays a crucial role in the cell spreading, cell migration and cholesterol-dependent Salmonella invasion [39]. *RILP* encodes a lysosomal protein related to the phagosomes in macrophages [40]. *SERPINF2* encodes serine protease inhibitor. The *SERPINF2* gene was involved in the recruitment of lymphocytes in the peripheral tissues and contributed to the maintenance of immunological functions that are related to IgE in human [41].

For IgG2, totally 7 candidate genes within 65.8 ~ 70.9 Mb on BTA21 were obtained. *AKT1*, encodes one member of AKT serine-threonine protein kinase family. Studies showed that Akt have important functions on immunity, including protecting thymocytes from cell death during the beta-selection checkpoint, affecting

peripheral B-cell maturation and survival and regulating the balance of mature B-cell [42–44]. *BCL11B* encodes a C2H2-type zinc finger protein, that a key regulator in the initial stages of human T-cell differentiation [45]. Previous studies discovered that BCL11B induced cell cycle arrest and protected against DNA damage-driven apoptosis [46, 47]. The protein encoded by *HHIPL1* is a secreted protein belonging to the glucose/sorbosone dehydrogenase family. In vitro studies revealed that *HHIPL1* enhanced hedgehog signaling and regulated cell proliferation and migration in human [48]. *DYNC1H1* encodes a member of the cytoplasmic dynein heavy chain family. KEGG enrichment showed *DYNC1H1* was involved in Salmonella infection pathway. The protein encoded by *HSP90AA1* is an inducible molecular chaperone that was upregulated and released from pancreatic β cells under pro-inflammatory conditions [49, 50]. The protein encoded by *TRAF3* is a member of the TNF receptor associated factor (TRAF) protein family. TRAF3 has been implicated as a negative regulator of NF-κB pathway activation and activator of antiviral type I IFN response to other DNA viruses. It can inhibit cell growth, colony formation and migration and resist to TNFα and cisplatin-induced cell death [51]. TRAF3 was also known to function as a resident nuclear protein to impact B cell metabolism and exert powerful restraint upon B cell survival and activation in human and mice [52, 53]. *KLC1* encodes kinesin-1 protein that is a member of the kinesin light chain family. kinesin-1 uncoated viral DNA and compromised nuclear pore complex (NPC) integrity, allowing viral genomes nuclear access to promote infection [54].

For IgA, 7 candidate genes were finally obtained. *IL6*, encodes a cytokine that has important functions in immune response, hematopoiesis, inflammation and the maturation of B cells [55]. This protein was primarily produced at sites of acute and chronic inflammation, where it was secreted into the serum and induced a transcriptional inflammatory response through interleukin 6 receptor, alpha [56, 57]. *PYCARD*, encodes an apoptosis speck-like protein (ASC), an adaptor protein of inflammasome complex that can enhance virus-induced type I interferon production with consequent reduction of virus replication to protect host upon virus infection [58]. *ITGAM* gene encodes the integrin alpha M chain belonging to leukocyte integrin Mac-1 that has been shown to mediate numerous effects of platelet factor 4 (PF4) on myeloid leukocytes and immune-modulation [59]. *TGFB1I1* encodes a coactivator of the androgen receptor and the deficiency of TGFB1I1 can protect chronic pancreatitis of human via down-regulation of the NF-κB (p65)/IL-6 signaling pathway [60]. *GUSB* encodes a hydrolase that degrades glycosaminoglycans and plays role in epatitis C virus (HCV)-infected human liver [61]. *CRCP* encodes a membrane protein of the CGRP family that was found to induce placental cellular differentiation, proliferation, and critical hormone signaling. Moreover, aberrant signaling of these CGRP family peptides during gestation associated with pregnancy disorders, like immunological effects, spontaneous abortion, preeclampsia and intrauterine growth restriction [62]. The protein encoded by *RABGEF1* forms a complex with *RABGEF1* that is required for endocytic membrane fusion. Millrine *et al.* identified Rabex-5, as an immunomodulatory drugs (IMiDs), targeted molecule to restrain toll-like receptor (TLR) and activated auto-immune promoting pathways [63]. *SBDS* encodes a highly conserved protein that plays an essential role in ribosome biogenesis. *SBDS* was identified as one of differentially expressed genes of immune genes between tumors and normal tissues, implying it might regulate tumor progression and migration [64].

The present study suggested 21 candidate genes for the concentrations of immunoglobulins in colostrum and serum in dairy cattle. Further in-depth investigations is needed to better understand the molecular

mechanism on how these genes regulated and impacted the formation of immunoglobulins in colostrum.

Conclusion

In this study, we conducted genome wide association studies for the concentrations of immunoglobulins in colostrum and serum in Chinese Holstein. A total of 25 genome-wide and 78 suggestive significant SNPs were detected for the total IgG, IgG1, IgG2, IgA and IgM traits, in which main quantitative trait loci for IgG traits were located in BTA19 and 21 and that for IgA were located in BTA25. Combining the identified significant SNPs, functional enrichment and the known QTL data, we identified 21 promising candidate genes for the concentrations of immunoglobulins in colostrum and serum, including *ABR, TIMM22, CRK, MYO1C, RILP, SERPINF2, AKT1, BCL11B, HHIPL1, DYNC1H1, HSP90AA1, TRAF3* and *KLC1* for IgG, and *IL6, PYCARD, ITGAM, TGFB1I1, GUSB, CRCP, RABGEF1* and *SBDS* for IgA. Our findings provided new insights into the genetic architectures underlying immunoglobulins concentrations in colostrum and important molecular information for the genetic improvement program on these traits in dairy cattle.

Abbreviations

Immunoglobulin: Ig; GWAS: genome-wide association study; SNP: Single nucleotide polymorphisms; RUNX3: RUNX family transcription factor 3; BTA: Bos taurus autosome; MMRA: Mixed model based single locus regression analyses; GRM: genetic relationship matrix; REML: restricted maximum likelihood; LD: linkage disequilibrium; QQ: Quantile-quantile; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; QTL: Quantitative trait locus; *ABR*: ABR activator of RhoGEF and GTPase; *TIMM22*: translocase of inner mitochondrial membrane 22; *CRK*: CRK proto-oncogene; adaptor protein; *MYO1C*:myosin IC; *RILP*: Rab interacting lysosomal protein; *SERPINF2*: serpin family F member 2; *AKT1*: AKT serine/threonine kinase 1; *BCL11B*: BAF chromatin remodeling complex subunit BCL11B; *HHIPL1*: HHIP like 1; *DYNC1H1*: dynein cytoplasmic 1 heavy chain 1; *HSP90AA1*: heat shock protein 90 alpha family class A member 1; *TRAF3*: TNF receptor associated factor 3; *KLC1:* kinesin light chain 1; *IL6*: interleukin 6; *PYCARD*: PYD and CARD domain containing; *ITGAM*: integrin subunit alpha M; *TGFB1I1*: transforming growth factor beta 1 induced transcript 1; *GUSB*: glucuronidase beta; *CRCP*: CGRP receptor component ; *RABGEF1*: RAB guanine nucleotide exchange factor 1; *SBDS*: SBDS ribosome maturation factor; GVHD: graft versus-host disease; TRAF: TNF receptor associated factor; NPC: nuclear pore complex; ASC: apoptosis speck-like protein; PF4: platelet factor 4; HCV: hepatitis C virus; IMiDs: immunomodulatory drugs; TLR: toll-like receptor.

Declarations

Acknowledgements

We appreciate the Beijing Dairy Cattle Center and the Beijing Sunlon Livestock Development Company Limited for providing the colostrum, serum and hair samples, and pedigree data for the Chinese Holstein cows used in this study.

Funding

This work was financially supported by the National Natural Science Foundation of China (31530070, 31872330, 31802041), Beijing Science and Technology Program (20200105, D171100002417001), and the Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62).

Authors' contributions

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

Availability of data and materials

All supporting data can be found within the additional files.

Ethics approval and consent to participate

All protocols for collection of the colostrum, 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. Samples collection specifically for this study followed the standard procedures with the full agreement of the Beijing Dairy Cattle Center and the Beijing Sunlon Livestock Development Company Limited who owned the animals

Consent for Publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Godden S. Colostrum management for dairy calves. Vet Clin N Am-Food A. 2008;24(1):19-+.
2. McGuirk SM, Collins M. Managing the production, storage, and delivery of colostrum. Vet Clin N Am-Food A. 2004;20(3):593-+.
3. Porter P. Immunoglobulins in bovine mammary secretions. Quantitative changes in early lactation and absorption by the neonatal calf. Immunology. 1972;23(2):225–38.
4. 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. J Immunol. 1996;157(8):3317–22.
5. Korhonen H, Marnila P, Gill HS. Milk immunoglobulins and complement factors. Br J Nutr. 2000;84(Suppl 1):75–80.

6. 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.
7. Kacprazak-Bergman I. Sexual dimorphism of heritability of immunoglobulin levels. Ann Hum Biol. 1994;21(6):563–9.
8. 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.
9. Outschoorn IM, Talor MV, Burek CL, Hoffman WH, Rose NR. Heritability analysis of IgG4 antibodies in autoimmune thyroid disease. Autoimmunity. 2014;47(5):320–6.
10. 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):2746–52.
11. 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):5467–73.
12. 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):3965–72.
13. 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. Anim Genet. 2013;44(5):485–92.
14. 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).
15. 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.
16. 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.
17. 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 Genom. 2018;19(1):694.
18. 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):11092–103.
19. Newby TJ, Bourne J. The nature of the local immune system of the bovine mammary gland. J Immunol. 1977;118(2):461–5.
20. Sasaki M, Davis CL, Larson BL. Production and turnover of IgG1 and IgG2 immunoglobulins in the bovine around parturition. J Dairy Sci. 1976;59(12):2046–55.
21. Browning BL, Browning SR. A Unified Approach to Genotype Imputation and Haplotype-Phase Inference for Large Data Sets of Trios and Unrelated Individuals. Am J Hum Genet. 2009;84(2):210–23.

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):559–75.
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. J Dairy Sci. 2009;92(7):3431–6.
24. Patterson HD, Thompson R. Recovery of Inter-Block Information When Block Sizes Are Unequal. Biometrika. 1971;58(3):545-+.
25. Johnson RC, Nelson GW, Troyer JL, Lautenberger JA, Kessing BD, Winkler CA, O'Brien SJ. Accounting for multiple comparisons in a genome-wide association study (GWAS). BMC Genom. 2010;11:724.
26. International HapMap C. A haplotype map of the human genome. Nature. 2005;437(7063):1299–320.
27. 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.
28. 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).
29. Haldar T, Ghosh S. Effect of Population Stratification on False Positive Rates of Population-Based Association Analyses of Quantitative Traits. Ann Hum Genet. 2012;76:237–45.
30. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21(2):263–5.
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):2225–9.
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:W316–22.
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):1084–97.
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):459–63.
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):5742–50.
37. Chiusolo V, Jacquemin G, Yonca Bassoy E, Vinet L, Liguori L, Walch M, Kozjak-Pavlovic V, Martinvalet D. Granzyme B enters the mitochondria in a Sam50-, Tim22- and mtHsp70-dependent manner to induce

apoptosis. Cell Death Differ. 2017;24(4):747–58.

38. 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):1019–32.
39. Brandstaetter H, Kendrick-Jones J, Buss F. Myo1c regulates lipid raft recycling to control cell spreading, migration and Salmonella invasion. J Cell Sci. 2012;125(Pt 8):1991–2003.
40. 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):170–4.
41. 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.
42. 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):12105–10.
43. 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):4043–50.
44. Chen J, Limon JJ, Blanc C, Peng SL, Fruman DA. Foxo1 regulates marginal zone B-cell development. Eur J Immunol. 2010;40(7):1890–6.
45. 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):2503–14.
46. 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. Journal of Investigative Dermatology. 2017;137(7):1523–32.
47. 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. *Molecular and Cellular Biology* 2018, 38(5).
48. Aravani D, Morris GE, Jones PD, Tattersall HK, Karamanavi E, Kaiser MA, Kostogrys RB, Ghaderi Najafabadi M, Andrews SL, Nath M, et al. HHIPL1, a Gene at the 14q32 Coronary Artery Disease Locus, Positively Regulates Hedgehog Signaling and Promotes Atherosclerosis. Circulation. 2019;140(6):500–13.
49. Ocana GJ, Perez L, Guindon L, Deffit SN, Evans-Molina C, Thurmond DC, Blum JS. Inflammatory stress of pancreatic beta cells drives release of extracellular heat-shock protein 90alpha. Immunology. 2017;151(2):198–210.
50. Eizirik DL, Colli ML, Ortis F. The role of inflammation in insulitis and beta-cell loss in type 1 diabetes. Nat Rev Endocrinol. 2009;5(4):219–26.
51. 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. Can Res. 2018;78(16):4613–26.

52. Bishop GA, Stunz LL, Hostager BS. TRAF3 as a Multifaceted Regulator of B Lymphocyte Survival and Activation. *Frontiers in Immunology* 2018, 9.
53. 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):11996–2010.
54. Strunze S, Engelke MF, Wang IH, Puntener D, Boucke K, Schleich S, Way M, Schoenenberger P, Burckhardt CJ, Greber UF. Kinesin-1-mediated capsid disassembly and disruption of the nuclear pore complex promote virus infection. Cell Host Microbe. 2011;10(3):210–23.
55. Jones GW, Hill DG, Cardus A, Jones SA. IL-27: a double agent in the IL-6 family. Clin Exp Immunol. 2018;193(1):37–46.
56. 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):401–6.
57. Frempah 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.
58. 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.
59. 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):6869–82.
60. Chen H, Tan P, Qian B, Du Y, Wang A, Shi H, Huang Z, Huang S, Liang T, Fu W. Hic-5 deficiency protects cerulein-induced chronic pancreatitis via down-regulation of the NF-kappaB (p65)/IL-6 signalling pathway. J Cell Mol Med. 2020;24(2):1488–503.
61. 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):386–90.
62. 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;814:229–40.
63. 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):10625–30.
64. Zhou H, Zhang C, Li H, Chen L, Cheng X. A novel risk score system of immune genes associated with prognosis in endometrial cancer. Cancer Cell Int. 2020;20:240.

Figures

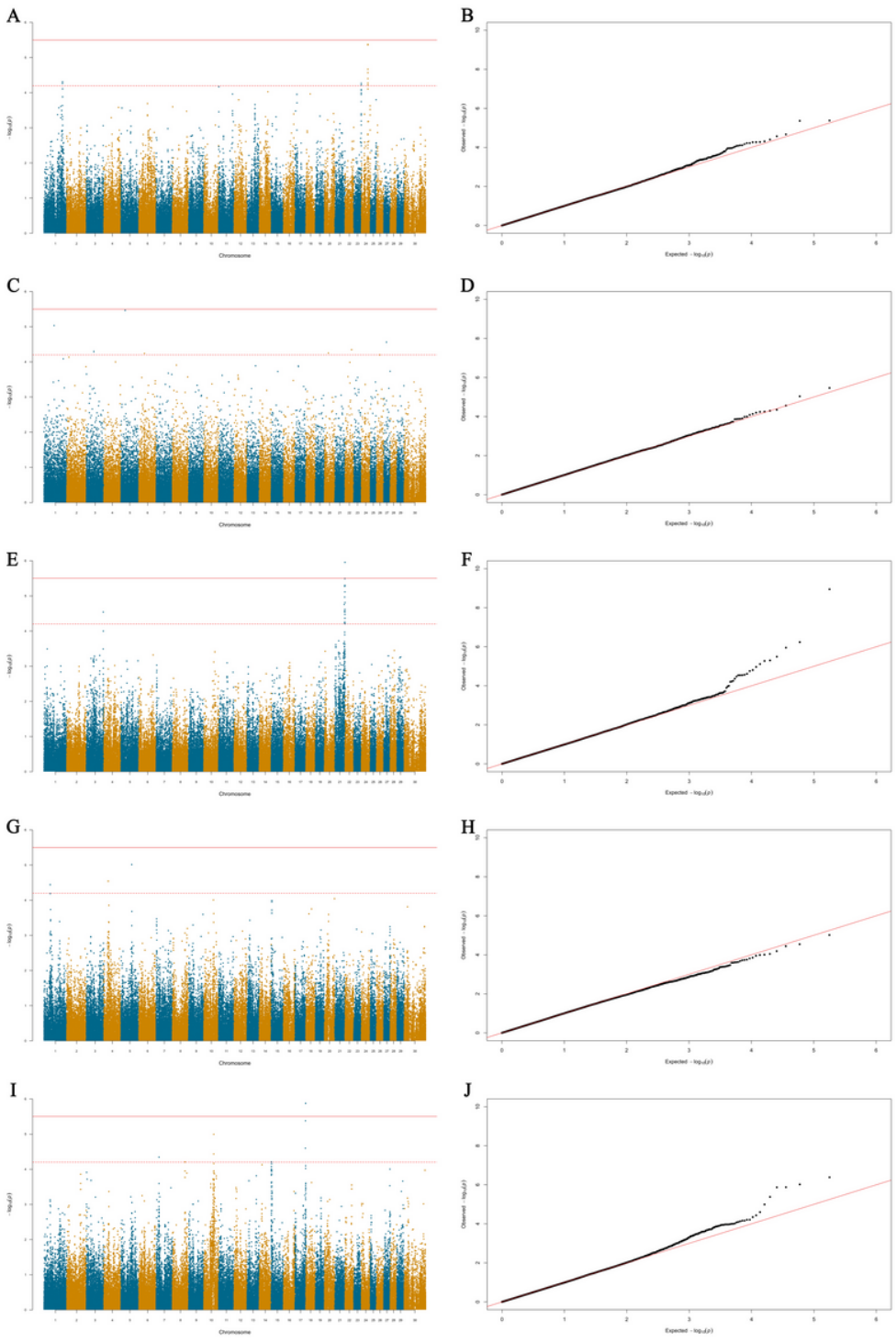


Figure 1

Manhattan and Q-Q plots of the observed P-values for the concentrations of immunoglobulins in the colostrum. (A) and (B) Indicated IgG concentrations. (C) and (D) Indicated IgG1 concentrations. (E) and (F) Indicated IgG2 concentrations. (G) and (H) Indicated IgA concentrations. (I) and (J) Indicated IgM concentrations. The Manhattan plots presented −log10 (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 indicated the genome-wide (3.16E−6) and suggestive significance (6.32E−5) thresholds,

respectively. The Q-Q plots showed the observed −log10-transformed P-values (y-axis) and the expected −log10-transformed P-values (x-axis).

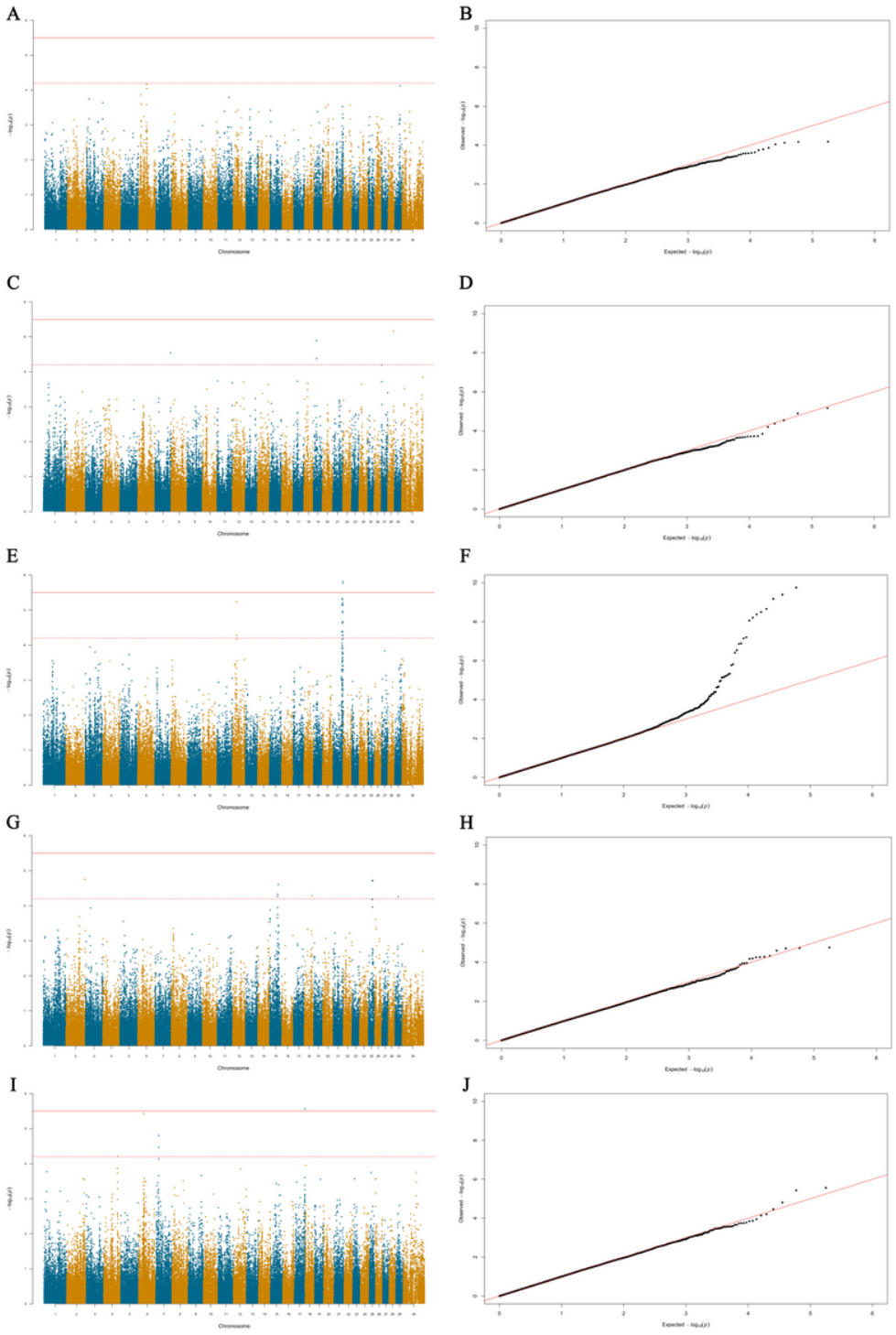


Figure 2

Manhattan and Q-Q plots of the observed P-values for the concentrations of immunoglobulins in the serum. (A) and (B) Indicated IgG concentrations. (C) and (D) Indicated IgG1 concentrations. (E) and (F) Indicated IgG2 concentrations. (G) and (H) Indicated IgA concentrations. (I) and (J) Indicated IgM concentrations.

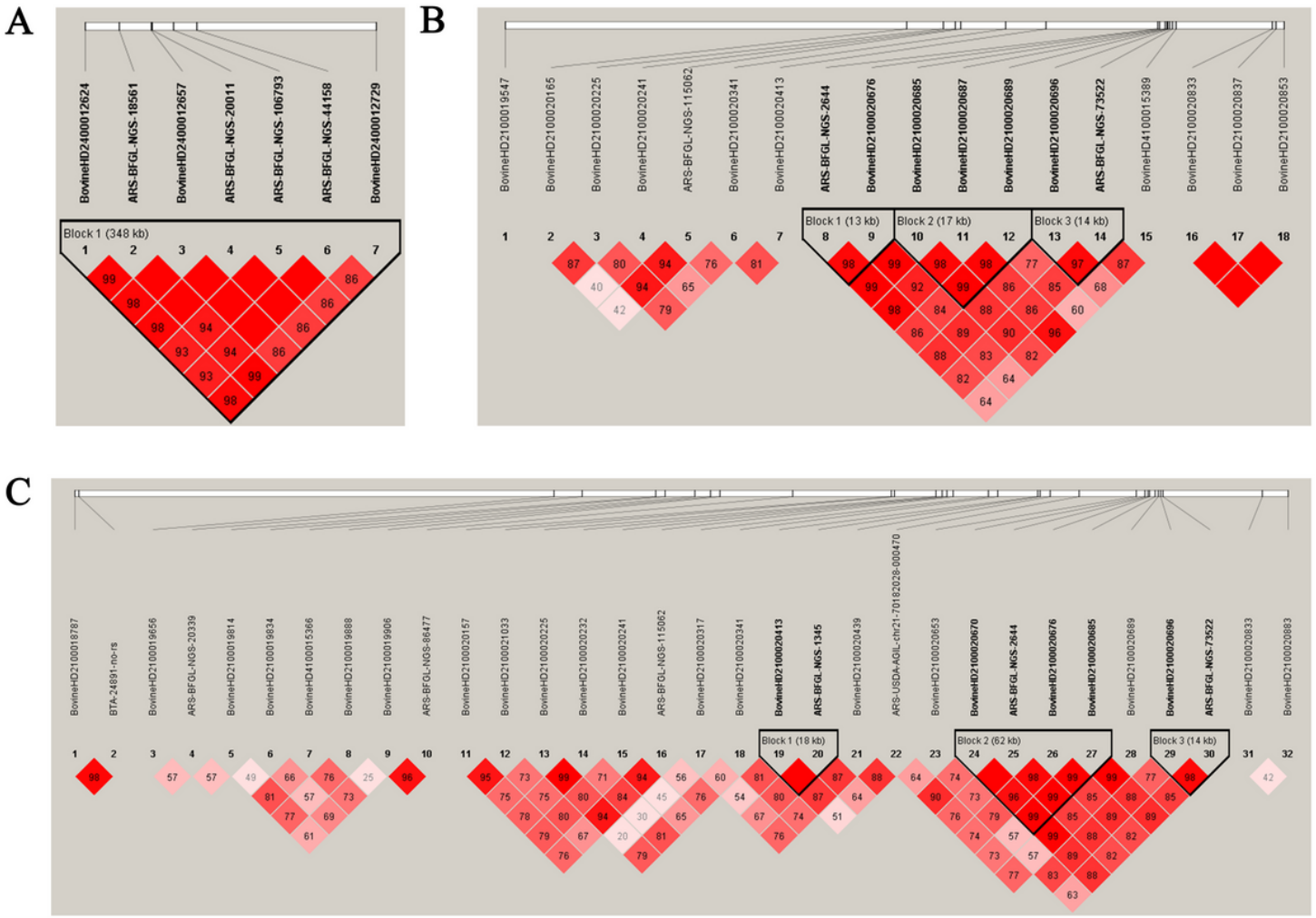


Figure 3

Haplotype blocks formed by the significant SNPs. (A) Indicated IgG concentration in colostrum. (B) Indicate IgG2 concentration in colostrum. (C) Indicate IgG2 concentration in serum. The black lines marked the identified blocks.

Supplementary Files

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

- AuthorChecklist.pdf
- TableS2.xlsx
- TableS1.xlsx
- PED.txt
- MAP.txt