ABSTRACT

In dairy cows, supernumerary teats (SNT) are not desired as they are considered a repository for bacteria; thus, SNT are a risk factor for mastitis. Supernumerary teats are a heritable oligo- or polygenic trait. The incidence of SNT in offspring must be reduced by genomic selection. However, in modern dairy farming, farmers often ignore the effects of SNT on cows. The study aimed to elucidate the effects of SNT on dairy cows from the blood transcriptome level and identify genes associated with SNT in Chinese Holstein cows. We selected 6 SNT cows (Yes) and 6 non-SNT cows (No). In the 6 SNT cows, 3 cows had 1 SNT (One) and 3 cows had 2 SNT (Two). They were divided into 3 comparison groups (One vs. No; Two vs. No; and Yes vs. No). RNA was extracted from blood white membrane cells of 12 cows, and RNA sequencing was performed. Differential gene expression analysis based on the negative binomial distribution was used to detect differentially expressed genes in the One versus No and Two versus No comparison groups. Genes that were significantly upregulated or downregulated both in the One versus No and Two versus No comparison groups. Genes that were significantly upregulated or downregulated both in the One versus No and Two versus No groups (shared genes, SG) were obtained for further analysis. We also performed gene set enrichment analysis for all genes expressed in the Yes versus No group, correlation analysis between SG and the hematological parameters, protein–protein interaction network analysis of SG to select hub genes, and alternative splicing analysis for Yes versus No group to explore the functions of differentially spliced genes. We detected 289 SG. Gene set enrichment analysis, gene ontology, and the Kyoto Encyclopedia of Genes and Genomes enrichment analysis results showed that SNT affect immunity, inflammation, and lactation-related pathways in dairy cows. Correlation analysis showed that LOC104968484, SLC25A6, GADD45G, BAX, APAF1, ATM, XIAP, MDM4, BDP1, CEP350, MED13, TAO1, SMG1, and RIF1 are associated with white blood cell count and absolute value of lymphocytes in SNT cows only, so they might be genes associated with SNT in Chinese Holstein cows. We found 2 genes (BAX and MDM4) were also differentially spliced genes. However, the causal relationship between these genes and the SNT phenotype needs to be further studied. This study is the first to reveal the adverse effects of SNT on dairy cows at a transcriptional level, and the genes we found can be used as a reference for further searching for candidate genes for the SNT phenotype.

Key words: supernumerary teats, Chinese Holstein, RNA sequencing

INTRODUCTION

Supernumerary teats (SNT) refers to teats in addition to normally developing functional teats (Pausch et al., 2012; Martin et al., 2016). In humans, SNT are associated with renal abnormalities and urinary diseases (pyuria, hematuria, and urolithiasis; Méhes, 1979; Hersh, 1988; Brown and Schwartz, 2004). In pigs, nonfunctional teats, which fail to produce milk and nurture young, are equivalent to SNT. The mammary glands of pigs are underdeveloped and cannot store milk, which is a limiting factor in piglet growth and survival (Hurley, 2019). Zhang et al. (2019) compared the mammary tissues of Hu sheep with or without SNT at different stages. During pregnancy and lactation, the mammary alveoli were denser in Hu sheep with SNT than in sheep without SNT (Zhang et al., 2019). The number of mammary alveoli is proportional to the lactation ability of the mammary gland (Brisken and Rajaram, 2006). Thus, Hu sheep with SNT may have higher milk yield than those without.
However, cows with SNT do not fare well. First, SNT not only have the potential to interfere with milking (Mackenzie and Marshall, 1925; Skjervold, 1960), but they are also considered a reservoir for bacteria; thus, SNT are a risk factor for mastitis (Joerg et al., 2014; Butty et al., 2017). In addition, SNT affect the milk performance of cows. Wen et al. (2019) compared the milk performance of 325 SNT cows and 737 non-SNT cows; specifically, their results showed that the daily milk yield and peak milk of SNT cows are lower than those of non-SNT cows, and the gap between them can reach 2 kg in summer (Wen et al., 2019). Yang et al. (2022) also showed that the daily milk yield, peak milk, and pre-milk yield of SNT cows are significantly lower than those of non-SNT cows (Yang et al., 2022). The existence of SNT affects the economic benefit of dairy farming. In modern dairy farming, SNT are mainly treated by artificial cutting. However, this method can increase the stress response of cattle and easily cause wound infections due to improper handling. Artificial cutting fails to meet the requirements of modern animal health and welfare (Fraser, 2008). Palacios and Abecia (2014) found that the milk yield of lactating sheep, whose SNT were cut off, significantly decreased compared with that of sheep with SNT (Palacios and Abecia, 2014). In cattle, the inheritance of SNT is oligo- or polygenic, with a heritability between 15 and 60% (Brka et al., 2002; Pausch et al., 2012). The heritability of SNT is different among breeds. For German Holsteins, the prevalence of SNT was 15%, and the posterior mean for the heritability was 0.09 with standard error of 0.08 (Brka et al., 2000). The incidence of SNT in Chinese Holstein cows was about 9.8%, and the heritability was about 0.22 in our previous study (Wen et al., 2021), which was medium heritability. Therefore, the incidence of SNT in cows must be reduced through genomic selection.

Studies on SNT mostly focused on Simmental and German Brown Swiss cows, and few studies worked on Holstein cows, which have stronger lactation performance than other cows (Brka et al., 2002). Functional genes associated with SNT were mainly identified by GWAS analysis (Pausch et al., 2012; Joerg et al., 2014; Butty et al., 2017). Our previous study identified 10 SNT-associated genes in BTA4 by single-step GWAS analysis (Wen et al., 2021). However, transcriptome sequencing (RNA-seq) has not been used to carry out relevant studies on cows with SNT. Blood is an important material in the process of disease research, and its composition changes represent the health status of animals. In addition, hematological parameters (HP) are easy to obtain. In view of these, blood RNA-seq was performed on Chinese Holstein dairy cows with and without SNT to explore the effects of SNT on cows and search for genes associated with SNT.

**MATERIALS AND METHODS**

**Ethics Statement**

The experimental procedures were approved by the Animal Welfare Committee of China Agricultural University, Beijing, China. Animal welfare and handling procedures were strictly followed throughout the experiment.

**Collection of Animals and Blood Samples**

The experimental cows came from a dairy farm of Beijing Shounong Group (Beijing, China). A total of 12 Chinese Holstein cows were selected based on their parity, lactation stage, and calving age. Half of the 12 cows were non-SNT cows (No), selected from non-SNT cows. The rest were SNT cows [Yes; including 3 cows with one SNT (One) and 3 cows with 2 SNT (Two)], selected from SNT cows. The experiment was divided into 3 comparison groups (One vs. No, Two vs. No, and Yes vs. No). About 20 mL of tail venous blood was collected from each cow. Ten milliliters of tail venous blood was used for RNA-seq, and the remaining 10 mL of tail venous blood was sent to Jinhaikeyu Company for HP testing (Beijing, China).

**mRNA Extraction and Sequencing**

Blood was immediately centrifuged at 1,057 × g for 15 min at 4°C to isolate white membrane layer and stored in liquid nitrogen for RNA-seq. We extracted total RNA from blood white membrane layer (mainly consists of platelets, lymphocytes, monocytes, and granulocytes) of 12 dairy cows by using TRIzol reagent (Invitrogen). The concentration and integrity of RNA were determined by NANO-DROP 2000 spectrophotometer and Bioanalyzer 2100 system (Agilent Technologies), respectively. The concentration of the RNA samples ranged from 150 ng/μL to 800 ng/μL, and their RNA integrity numbers were greater than 8 (Supplemental Table S1; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022), which was sufficient for gene expression analysis (Cánovas et al., 2014; Gallego Romero et al., 2014; Davila et al., 2016). Novogene Co. Ltd. (Beijing, China) performed the cDNA library preparation and RNA sequencing, generating paired-end reads at 150 bp length on the Illumina HiSeq 4000 platform.
RNA-seq Data Processing and Analysis

We conducted quality assessment for raw reads using FastQC (version 0.11.3; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Subsequently, we filtered adapters and reads with low quality by using the NGSQC.Toolkit software (version 2.3; https://github.com/mjain-lab/NGSQCToolkit/releases/tag/v2.3) to obtain clean reads. The proportion of bases with low quality in reads >30% were considered reads with low quality. Bases with low quality were those with a mass less than 20. The SAM files were observed after Hisat2 software (version 2.1.0; https://daehwankimlab.github.io/Hisat2/download/) was used to align with the bovine reference genome (http://ftp.ensembl.org/pub/release-97/gtf/bos_taurus/Bos_taurus.ARS-UCD1.2.97.gtf.gz). The details of sequencing data are shown in Supplemental Table S2 (https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022). The SAM files were then converted to BAM files by SAMtools (version 1.16.1; https://github.com/samtools/samtools/releases/tag/1.16.1). Rsubread (version 2.10.5; http://bioconductor.org/packages/release/bioc/html/Rsubread.html) software was used to generate the read counts for each sample. Genes that were not expressed in all samples were deleted, and read counts were corrected to fragments per kilobase of transcript per million mapped reads (Supplemental Table S3; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022). The DESeq2 package was used to find significantly differentially expressed genes (DEG) in the One versus No and Two versus No groups [log$_2$-fold change $>$ 1, adjusted $P$-value ($P$-adjust) $<$ 0.05]. Given that our study focused on stable SNT genetic markers of cows. The “marker” in our study implied significantly DEG shared by group One versus No and group Two versus No in the same direction of difference (shared genes, SG). In other words, SG were upregulated in both One versus No and Two versus No groups (up-SG) and significantly downregulated in both One versus No and Two versus No groups (down-SG). Therefore, further analysis of SG is meaningful. Gene set enrichment analysis (GSEA) was performed using all genes in the Yes versus No group by GSEAbase package (version 1.36.0; http://www.bioconductor.org/packages/3.4/bioc/html/GSEABase.html).

To further search for genes associated with SNT of Chinese Holstein cows, we performed correlation analysis between the expression level of SG and HP (the blood routine test data shown in Supplemental Table S4; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022) of cows using the lm function of R (version 4.2.1; https://cran.r-project.org/). To understand the relationship between proteins encoded by SG, protein–protein interaction (PPI) network analysis was performed via Cytoscape software (version 3.9.1; https://cytoscape.org/). CytoHubba within Cytoscape (version 3.9.1) was then used for hub gene analysis, and maximal clique centrality was used as the rank criteria. To carry out alternative splicing analysis, STAR software (version 2.7.9; https://github.com/alexdobin/STAR/releases/tag/2.7.9a) was used for genome assembly to obtain BAM files. The rMATS (version 4.1.2) software (http://rnaseq-mats.sourceforge.net/download.html) was used for alternative splicing analysis. Five types of splicing events were recognized by rMATS (version 4.1.2), namely, skipped exon, alternative 5′ splice site, alternative 3′ splice site, mutually exclusive exons, and retained intron. The rMATS default parameters were used for the differential splicing analysis of the Yes versus No group. Transcripts with $P$ $<$ 0.05 were considered differentially spliced transcripts. Intersections between differentially spliced genes (DSG) of different splicing events and SG were shown by UpSetR package (version 1.4.0; https://cran.rstudio.com/web/packages/UpSetR/). The results of DSG were visualized by rmats2sashimiplot software (version 2.0.4; https://github.com/Xinglab/rmats2sashimiplot/releases/tag/v2.0.4). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed for SG and DSG using the ClusterProfiler R package. Gene Ontology terms and KEGG pathways with $P$ $<$ 0.05 were obtained.

Validation of Sequencing Results by Real-Time Quantitative PCR

PrimeScript RT reagent Kit with gDNA Eraser (Takara Biotechnology Co., Ltd.) was used to produce cDNA from qualified RNA. The bovine gene GAPDH was used as the reference gene, and 12 DEG were randomly selected. Primers were designed through online website Primer-Blast, and the primers’ information is shown in Table 1. The primer sequences were synthesized by Tsingke Biotechnology Co., Ltd. (Beijing, China). Three technical replicates were prepared for each cDNA sample, and the mean cycle threshold (Ct) value was used to calculate the genes’ relative expression by the 2$^{-\Delta\Delta C_t}$ method. The significant differences between groups were examined by t-test.

RESULTS

Differential Expression Profile of Genes Between Cows With and Without SNT

A total of 397 and 1,559 DEG were detected in the One versus No and Two versus No groups, respectively
As shown in the clustering heatmap (Figure 1A), the difference in the expression patterns of DEG in the 2 comparison groups was obvious. In addition, a total of 20,303 genes expressed in the Yes versus No group were used for principal component analysis and hierarchical clustering. The results showed that SNT could classify the groups as left side and right side, which illustrated SNT as one of the main influencing factors of the differences in samples (Figure 1B). In DEG, 134 and 406 genes were significantly upregulated in the One versus No and Two versus No groups, respectively, and most of the DEG were significantly downregulated in the One versus No (263) and Two versus No (1,153) groups (Figure 1C). Among them, we found 84 genes that were significantly upregulated in the One versus No and Two versus No groups (left panel in Figure 1D), and 205 genes that were significantly downregulated in the One versus No and Two versus No groups (right panel in Figure 1D).

### Genes Expressed Between SNT Cows and Non-SNT Cows, or SG Enriched in Immunity or Inflammation-Related Pathways

First, we used all genes expressed in the Yes versus No group for GSEA. The results showed that the SNT phenotype was positively correlated with inflammation-related pathways ($P < 0.05$), such as systemic lupus erythematosus (SLE), neutrophil extracellular trap (NET) formation, necroptosis, p53 signaling pathway, NF-κB signaling pathway, IL-17 signaling pathway, *Staphylococcus aureus* infection, NOD-like receptor signaling pathway, and TNF signaling pathway (Figure 2A). Moreover, the SNT phenotype was negatively correlated with the prolactin signaling pathway and carbohydrate digestion and absorption (Figure 2A). Shared genes were selected for functional enrichment analysis. Gene Ontology and KEGG enrichment analyses were carried for SG. A total of 80 notable GO terms were enriched ($P < 0.05$; Supplemental Table S7; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022). The top 40 terms of biological process were chosen based on the $P$-value for visualizations (Figure 2B), and these terms were mainly associated with immune regulation (e.g., T cell cytokine production, T cell-mediated immunity, and regulation of immune response), heme metabolic process, and post-transcriptional gene silencing. Posttranscriptional gene silencing may be associated with downregulation of most genes in SNT cows. A total of 22 pathways were enriched by SG ($P < 0.05$; Supplemental Table S8; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022), which were mainly associated with immunity, inflammation, and apoptosis (e.g., SLE, NET formation, apoptosis, and p53 signaling pathway and NF-κB signaling pathway; Figure 2C).

### Table 1. The primer information for real-time quantitative PCR (F = forward, R = reverse)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence</th>
<th>Product length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAMDEC1</td>
<td>F: AGCCAAGATGCCTACTGCAA R: GCTTGGATTCATTGAGTTTGGA</td>
<td>237</td>
</tr>
<tr>
<td>ADIRF</td>
<td>F: GCCGCCATCCCTTAACCACCTC R: TGGGCTAGTCTTGGAACACTT</td>
<td>189</td>
</tr>
<tr>
<td>C1QC</td>
<td>F: GCCGCCATCCTGCTACAAAACCCTTCCATACCCA R: CTGCGAGTGCTGGGCAACACCA</td>
<td>140</td>
</tr>
<tr>
<td>CCL3</td>
<td>F: ACACCTGCTTCTCGGACAGCACTTGA R: ATGAGGTCGTTGCTGCCGACACACCA</td>
<td>291</td>
</tr>
<tr>
<td>CLEC1B</td>
<td>F: AGCTACGGTCCTGGGAGTGATT R: GCTGGCAGTCTTTCTCCATACACC</td>
<td>166</td>
</tr>
<tr>
<td>FSCN1</td>
<td>F: CAAGGAGCGCCAGATGCTACT R: ATGTTGTAGGCACCGTCTCCG</td>
<td>288</td>
</tr>
<tr>
<td>HECTD4</td>
<td>F: CTTTGAGGGAAGAGACATTGA R: ATGAGGGCAGTTGTTTCGCT</td>
<td>124</td>
</tr>
<tr>
<td>LOC10496834</td>
<td>F: CCTACTCCTTGTTCTTCTTGACAGGG R: TGAATGCTCCTCACAGAACACCC</td>
<td>218</td>
</tr>
<tr>
<td>LOC618541</td>
<td>F: GCTCTGACTGCGAGCTCTCA R: GACCTGTGAACACCAACTACGG</td>
<td>222</td>
</tr>
<tr>
<td>PLAC5A</td>
<td>F: ACCAGTTGTGTCTTCACAGCCA R: GTCTCAGGGCGCACACCTGGA</td>
<td>193</td>
</tr>
<tr>
<td>PRSS57</td>
<td>F: CATCACGGACCCCACTACCC R: CGATTCTCAGGAAGTCCGGCA</td>
<td>190</td>
</tr>
<tr>
<td>RETN</td>
<td>F: GCAGGTCCTTAAAAAAGGGAGCG R: TCCCTACTGCCCCAGGAAGAAT</td>
<td>222</td>
</tr>
<tr>
<td>GAPDH</td>
<td>F: GGTCCTGAGTATGTTGCTGGGGA R: GCCATTGCCTGACAAATCTTGA</td>
<td>119</td>
</tr>
</tbody>
</table>

($\log_2$-fold change) > 1, $P$-adjust < 0.05; Supplemental Tables S5, S6; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022).
ated with immunity and inflammation of dairy cows. Supernumerary teats may have adverse effects on the health of cows and affect lactation performance.

**Strong Association Between Genes and HP in SNT Cows Only**

To further search for genes associated with SNT, correlation analysis between SG and HP was conducted. We found that LOC104968484, SLC25A6, GADD45G, and BAX were significantly, or tended to be significantly, negatively correlated with white blood cell (WBC) count and absolute value of lymphocytes (W-SCC) only in SNT cows (green line in Figure 3A), and they were significantly upregulated in SNT cows (Figure 2C). By contrast, the expression levels of APAF1, ATM, XIAP, and MDM4, which were significantly downregulated in SNT cows (Figure 2C), were significantly, or tended to be significantly, positively correlated with WBC count and W-SCC in SNT cows (green line in Figure 3B). However, the upregulated genes (LOC104968484, SLC25A6, GADD45G, and BAX)

Figure 1. Differentially expressed genes (DEG) between cows with supernumerary teats (SNT) and non-SNT cows. (A) Heatmaps of DEG in the One versus No and Two versus No groups. (B) Principal component analysis (PCA) of 20,303 genes expressed in SNT and non-SNT cows. (C) Volcano plot of DEG in the One versus No and Two versus No groups. (D) Genes that were significantly upregulated both in the Two versus No and One versus No groups, and significantly downregulated both in the Two versus No and One versus No groups. Up = significantly upregulated genes; down = significantly downregulated genes; No = non-SNT cows; Yes = SNT cows; One = cows with 1 SNT; Two = cows with 2 SNT; up-SG = genes that were significantly upregulated both in the One versus No and Two versus No groups; down-SG = genes that were significantly downregulated both in the One versus No and Two versus No groups. FDR = false discovery rate.
Figure 2. Functional enrichment analysis of genes. (A) Gene set enrichment analysis was performed for 20,303 genes expressed in the Yes versus No group. The supernumerary teat (SNT) phenotype was significantly positively correlated with inflammation-related pathways and significantly negatively correlated with lactation-related pathways. The terms highlighted in red are those associated with immunity. (C) The left shows the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched by SG, whereas the genes shown on the right are the genes in the corresponding pathway. These genes were chosen based on correlation analysis, which found that these genes were strongly correlated with white blood cell count and absolute value of lymphocytes only in SNT cows. The y-axis is corrected read counts of genes. No = non-SNT cows; Yes = SNT cows.
Figure 3. Correlation between the expression level of important genes associated with supernumerary teats (IGR-SNT) and (a) white blood cell (WBC) count and (b) absolute value of lymphocytes (W-SCC) in cows with supernumerary teats (SNT) and non-SNT cows, respectively. (A) Correlation between the expression level of upregulated IGR-SNT (LOC104968484, SLC25A6, GADD45G, and BAX) and WBC count and W-SCC. The upregulated IGR-SNT were significantly negatively correlated \( P < 0.05 \) or tended to be negatively correlated with WBC count and W-SCC only in SNT cows. (B) Correlation between the expression level of downregulated IGR-SNT (APAF1, ATM, XIAP, and MDM4) and WBC count and W-SCC. The downregulated IGR-SNT were significantly positively correlated \( P < 0.05 \) or tended to be positively correlated with WBC count and W-SCC only in SNT cows. The x-axis is the corrected read counts of IGR-SNT.
Figure 4. Correlation between the expression level of important genes associated with supernumerary teats (IGR-SNT) and (m) red blood cell (RBC) count and (n) platelet (PLT) count in cows with supernumerary teats (SNT) and non-SNT cows, respectively. (A) Correlation between the expression level of upregulated IGR-SNT (LOC104968484, SLC25A6, GADD45G, and BAX) and RBC count and PLT count. The upregulated IGR-SNT had no correlation with RBC count and PLT count in SNT cows or non-SNT cows. (B) Correlation between the expression level of downregulated IGR-SNT (APAF1, ATM, XIAP, and MDM4) and RBC count and PLT count. The downregulated IGR-SNT had no correlation with RBC count and PLT count in SNT cows or non-SNT cows. The x-axis is the corrected read counts of IGR-SNT.
Figure 5. Protein–protein interaction (PPI) network analysis of shared genes (SG). (A) PPI network of SG. Green nodes = significantly downregulated genes in cows with supernumerary teats (SNT); red nodes = significantly upregulated genes in SNT cows. Nodes with black circles were important genes associated with SNT (IGR-SNT); nodes with red circles were hub genes. (B) The top 10 hub genes ranked by the maximal clique centrality method, which was in CytoHubba software (version 2.7.9).
and downregulated (APAF1, ATM, XIAP, and MDM4) genes did not have such a relationship with WBC count and W-SCC in non-SNT cows (red line in Figures 3A and 3B). More interestingly, any significant correlation was not observed between the 8 genes (APAF1, ATM, XIAP, MDM4, LOC104968484, SLC25A6, GADD45G, BAX, APAF1, ATM, XIAP, and MDM4) in SNT cows. Plus sign = positive correlation; minus sign = negative correlation; arrow pointing down = downregulated genes in SNT cows; arrow pointing up = upregulated genes in SNT cows; WBC = white blood cell; W-SCC = absolute value of lymphocytes.

Figure 6. The correlation between the expression level of the 6 hub genes and the important genes associated with supernumerary teats (IGR-SNT). (A) Correlation between the expression level of hub genes (BDP1, CEP350, MED13, TAOK1, SMG1, and RIF1) and downregulated IGR-SNT (APAF1, ATM, XIAP, and MDM4) in cows with supernumerary teats (SNT). The y-axis shows the corrected read counts of downregulated IGR-SNT. The x-axis shows corrected read counts of hub genes. (B) Schematic of the relationship between the 6 hub genes (BDP1, CEP350, MED13, TAOK1, SMG1, and RIF1) and IGR-SNT (LOC104968484, SLC25A6, GADD45G, BAX, APAF1, ATM, XIAP, and MDM4) in SNT cows. Plus sign = positive correlation; minus sign = negative correlation; arrow pointing down = downregulated genes in SNT cows; arrow pointing up = upregulated genes in SNT cows; WBC = white blood cell; W-SCC = absolute value of lymphocytes.

CytoHubba was then used for screening hub genes, and the rank criteria was selected using the maximal clique centrality method. In our study, the top 10 genes were considered hub genes, and a high degree of interaction was found between the 8 genes (APAF1, ATM, XIAP, MDM4, LOC104968484, SLC25A6, GADD45G, BAX) and red blood cell count and platelet count in SNT or non-SNT (Figures 4A and 4B). The results suggested that the 8 genes (APAF1, ATM, XIAP, MDM4, LOC104968484, SLC25A6, GADD45G, and BAX) were highly associated with immunity of SNT cows, and they could be considered important genes associated with SNT (IGR-SNT).

We conducted PPI network analysis for SG, and the results showed that highly interacting proteins, including SMG1, BDP1, CEP350, MED13, AFF4, PHIP, TAOK1, ATM, MAP3K2, PIK3C2A, TAOK1, BRWD3, and ESR1, were encoded by genes that were significantly downregulated in SNT cows (Figure 5A).
SNT cows only. They could be considered genes associated with SNT. The differences between IGR-SNT and the 6 hub genes were that in SNT cows, the r values (correlation coefficients) of correlation between IGR-SNT, WBC count, and W-SCC were generally higher than that of the 6 hub genes (right panel in Supplemental Figure S1C; Chen, 2022). The P-value showed the opposite trend (left panel in Supplemental Figure S1C; Chen, 2022). In addition, based on the result of KEGG enrichment analysis for SG, the 6 hub genes (BDP1, CEP350, MED13, TAOK1, SMG1, and RIF1) were not enriched in any prominent pathways, but the IGR-SNT were all enriched (Figure 2C; Supplemental Table S8; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022).

**Hub Genes of SG were Highly Correlated with IGR-SNT**

Hub genes play an important role in the PPI network. In other words, the regulation of other genes is often influenced by the hub gene. Further analysis revealed that the expression levels of the 6 hub genes (BDP1, CEP350, MED13, TAOK1, SMG1, and RIF1) in SNT cows were almost significantly positively correlated with the downregulated IGR-SNT (APAF1, ATM, XIAP, and MDM4; Figure 6A). In SNT cows, although the P-value of correlation between the 6 hub genes and upregulated IGR-SNT (LOC104968484, SLC25A6, GADD45G, and BAX) was almost above 0.05, they all showed a negative correlation trend (Supplemental Figure S2; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022). Thus, a strong regulatory relationship was found between the 6 hub genes and the IGR-SNT, and the relationship between the 6 hub genes and the IGR-SNT is shown in Figure 6B.

**Alternative Splicing Analysis of Genes Between SNT Cows and Non-SNT Cows**

To explore the mechanism of alternative splicing in the formation of SNT, rMATS were used to examine DSG of different splicing events. A total of 3,060 DSG were obtained in the Yes versus No group (P < 0.05). More than half of the DSG splicing type were skipped exon. The results of differentially spliced transcripts of different splicing events are shown in Supplemental Tables S9 through S13 (https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022). A total of 57 SG were DSG (Figure 7A). A KEGG enrichment analysis was carried out for DSG. A total of 79 pathways were enriched (P-adjust < 0.05; Supplemental Table S14; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022). The top 30 pathways were mainly associated with immunity and inflammation (Figure 7B). More importantly, 2 IGR-SNT (BAX and MDM4) were identified as DSG (Figure 7C and D).

**Validation of RNA-seq Results by Real-Time Quantitative PCR**

To verify the results of DEG obtained by RNA-seq, we randomly selected 12 genes from the One versus No and Two versus No groups to perform real-time quantitative PCR. The comparison results of RNA-seq and real-time quantitative PCR showed the high reliability of RNA-seq (Supplemental Figure S3; https://doi.org/10.6084/m9.figshare.20186543.v6; Chen, 2022).

**DISCUSSION**

This work is the first study to reveal the adverse effects of SNT on dairy cows at transcriptional level and explore the genes associated with SNT in Chinese Holstein cows by using RNA-seq technology. Inflammation-related pathways were activated, and prolactin signaling pathway and carbohydrate digestion and absorption were inhibited in SNT cows. The GSEA results showed that the SNT phenotype was positively correlated with inflammation-related pathways (e.g., IL-17 signaling pathway. Staphylococcus aureus infection, TNF signaling pathway, and NF-κB signaling pathway), and negatively correlated with prolactin signaling pathway and carbohydrate digestion and absorption. Staphylococcus aureus is an important gram-positive bacterial pathogen that causes IMI, and is associated with various forms of subclinical and clinical mastitis in dairy cows (Liu et al., 2022). Supernumerary teats may constitute a minor, but notable, risk factor for mastitis (Vasileiou et al., 2019), which was consistent with our results, which found that the S. aureus infection pathway was activated in SNT cows. Studies have shown that SNT cows have poor milk performance (Wen et al., 2019; Yang et al., 2022). Our result found that the prolactin signaling pathway and carbohydrate digestion and absorption were inhibited in SNT cows. Prolactin plays an important role in lactation, which can induce breast acinar cells to produce milk protein (Zhang et al., 2021). Long-term inhibition of prolactin reduces milk yield of dairy cows (Lacasse et al., 2011; Lollivier et al., 2015). Carbohydrate is the main component of ruminant diet and can be divided into fibrous carbohydrate and NFC. Non-fibrous carbohydrate is an important dietary energy source for ruminants, which can affect the performance of ruminants (Batajoo and Shaver, 1994; Gao and Oba, 2016). Batajoo and Shaver (1994) found that decreasing dietary NFC decreased DMI, milk protein percentage, and production in dairy...
Figure 7. Alternative splicing analysis of genes between cows with supernumerary teats (SNT) and non-SNT cows. (A) Intersection between differentially spliced genes (DSG) of different splicing events and shared genes (SG). (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of DSG (top 30). The pathways highlighted in red were associated with immunity or inflammation. (C) Alternative splicing visualization of the 2 genes (BAX and MDM4), which were both differentially expressed and differentially spliced genes. A3SS = alternative 3′ splice site; A5SS = alternative 5′ splice site; RI = retained intron; MXE = mutually exclusive exons; SE = skipped exons.
cows. Thus, we speculated that the decreased lactation performance of SNT cows in other studies (Wen et al., 2019; Yang et al., 2022) may be partly associated with the inhibition of prolactin and carbohydrate digestion and absorption in cows with SNT. Shared genes were used for functional enrichment analysis. Multiple immunity-related GO terms and inflammation-related pathways were enriched by SG, such as NET formation. Neutrophil extracellular traps are extracellular DNA fibers composed of histones and neutrophil antimicrobial proteins; particularly, they are released from neutrophils to kill pathogens (Klopf et al., 2021). Although NET play an important role in host defense under infection, NET formation has recently been implicated in the pathophysiology of autoimmunity (Grayson and Kaplan, 2016). For example, extensive NET formation has been shown to aggravate SLE (Villanueva et al., 2011). In our study, SLE was the most prominent pathway in KEGG enrichment. Furthermore, GSEA results revealed that NET formation and SLE were positively correlated with the SNT phenotype. Pausch et al. (2012) and Peng et al. (2017) both detected the Wnt signaling pathway, which plays an important role in the formation of SNT through GWAS analysis (Pausch et al., 2012; Peng et al., 2017). However, the Wnt pathway was not enriched by RNA-seq in our study.

In correlation analysis, we found genes (LOC104968484, SLC25A6, GADD45G, BAX, XIAP, APAFI, ATM, MDM4, BDP1, CEP350, MED13, TAOK1, SMG1, and RIF1) that had a strong correlation with WBC count and W-SCC only in SNT cows. The relationship between these genes and blood routine test parameters suggested that the SNT phenotype affected the immunity of cows. The LOC104968484, SLC25A6, GADD45G, and BAX genes were negatively correlated with WBC count and W-SCC only in SNT cows, and the expression levels of these genes were significantly upregulated in SNT cows. Immunoglobulin heavy variable 4-59 (LOC104968484) and solute carrier family 25 member 6 (SLC25A6) were enriched in the ENT formation pathway. Growth arrest and DNA damage 45G (GADD45G) and B-cell lymphoma-2-associated X protein (BAX) promote apoptosis (Edlich, 2018; Shin et al., 2019). Apoptosis has been described as a potential mechanism involved in tissue destruction, which is caused by infectious and inflammatory diseases, such as mastitis (Mi et al., 2021). Raj et al. (2022) showed that the BAX gene has an adverse effect on milk yield (Raj et al., 2022). After heat stress, the relative survival rate of mammary epithelial cells decreases significantly, and the expression level of BAX increases significantly (Yang et al., 2022). Therefore, the upregulation of BAX in dairy cows may predict adverse results.

Additionally, 4 genes (XIAP, APAFI, ATM, and MDM4) were positively correlated with WBC count and W-SCC only in SNT cows, and these genes were downregulated in SNT cows. X-linked inhibitor of apoptosis protein (XIAP) is a crucial anti-apoptotic protein, which is the only inhibitor of the apoptosis protein to directly interact with caspases (Shang et al., 2019). Studies have shown that deficiency of the XIAP protein in humans leads to immunodeficiency and inflammatory bowel disease (Rigaud et al., 2006). Activation of T cells plays an important role in the clearance of foreign antigens. When antigens are removed, the activated T cells should be removed to avoid unnecessary inflammation. Apoptosomes in T cells play an important role in this process (Reiner, 2009), and apoptotic peptidase activating factor 1 (APAF1) participates in the formation and activation of the apoptosome (Zou et al., 1997). However, APAFI was downregulated in SNT cows, which indicated that SNT cows were likely to suffer from unnecessary inflammation. ATM affects mammary duct development and mammary epithelial hyperplasia (Bowen et al., 2005). Mice without ATM failed to develop mature mammary glands (Bowen et al., 2005; Dyer et al., 2017), indicating the importance of ATM to mammary gland development. MDM4 is an inhibitor of p53 (Kawai et al., 2007), and the p53 protein induces apoptosis (Wawryk-Gawda et al., 2014). These results suggested that SNT may affect the apoptosis and immunity of dairy cows. Numerous studies have shown that apoptosis, inflammation, and immunity exist almost simultaneously in most cases (Chen et al., 2022; Ma et al., 2022; Meng et al., 2022).

In PPI network analysis, we found 6 hub genes (BDP1, CEP350, MED13, TAOK1, SMG1, and RIF1) are also associated with SNT, and a close relationship was observed between the 6 hub genes and the IGR-SNT (LOC104968484, SLC25A6, GADD45G, BAX, XIAP, APAFI, ATM, and MDM4). The 6 hub genes were significantly, or tended to be significantly, positively correlated with WBC count and W-SCC only in SNT cows. BDP1, the hub gene with the highest score, is a subunit of transcription initiation factor IIIB (TFIIIB). Initiation of gene transcription by RNA polymerase III (Pol III) requires the activity of TFIIIB (Gouge et al., 2017). In addition to BDP1, CEP350 (Kumar et al., 2013) and MED13 (Zhou et al., 2021) contribute to the cell cycle. The TAOK1 protein negatively regulates IL-17-mediated signaling and inflammation (Zhang et al., 2018). The SMG1 protein mediates post-transcriptional regulation of genes (Lloyd and Davies, 2013; Roberts et al., 2013). PIF1 is involved in DNA replication (Blasiak et al., 2021). The 6 hub genes that were downregulated in SNT cows were positively correlated with downregu-
lated IGR-SNT and negatively correlated with upregulated IGR-SNT. Thus, a strong regulatory relationship exists between the hub genes of SG and IGR-SNT. The expression levels of IGR-SNT and the 6 hub genes all have a correlation with WBC count and W-SCC only in SNT cows. Therefore, these genes are associated with the SNT phenotype. However, we do not know the causal relationship between these genes and the SNT phenotype. That is, the presented link between these genes and SNT may be due to a higher infection load in SNT cows (SNT holding the bacteria). It is also possible that these genes are the cause of the SNT phenotype, but we did not conduct trials to prove this in the current study, which is a limitation of this study.

Alternative splicing may mediate the effect of SNT on dairy cows and the formation of SNT. Alternative splicing, which plays an active role in normal growth and development of higher eukaryotes, contributes to cell differentiation, organ development, and tissue homeostasis (Wang et al., 2008). In the alternative splicing analysis, about one-fifth of SG were DSG, and 2 IGR-SNT (BAX and MDM4) were DSG. Thus, alternative splicing may mediate the effect of SNT on dairy cows or formation of SNT. The KEGG enrichment analysis of DSG also showed multiple inflammatory and immune-related pathways that were significantly enriched, such as the T cell receptor pathway, NOD-like receptor signaling pathway, MAPK signaling pathway, and chemokine signaling pathway. The mechanism of alternative splicing participating in the effect of SNT on dairy cows needs to be further studied.

CONCLUSIONS

Our study is the first to demonstrate that the SNT phenotype is positively associated with inflammation-related pathways, and negatively associated with the prolactin signaling pathway and carbohydrate digestion and absorption from the transcriptional level, by comparing blood transcriptome sequencing data of SNT and non-SNT cows. We screened genes (LOC104968484, SLC25A6, GADD45G, BAX, APAF1, ATM, XIAP, MDM4, BDP1, CEP350, MED13, TAOK1, SMG1, and RIF1) associated with SNT through correlation analysis. The pattern of correlation between genes (IGR-SNT and the 6 hub genes) and WBC count and W-SCC showed that the SNT phenotype will reduce the immunity of cows. Therefore, the SNT phenotype has an adverse effect on dairy cows, including decreased immune response and increased inflammation, and may affect lactation. The genes we found can be used as a reference for further searching for candidate genes of the SNT phenotype.

ACKNOWLEDGMENTS

This work was funded by the National Key R&D Program of China (grant nos. 2021YFD1200903 and 2021YFD1200900, Beijing, China), the NSFC-PSF joint project (grant no. 31961143009, Beijing, China), the Beijing Dairy Industry Innovation Team (grant no. BAIC06, Beijing, China), the Beijing Natural Science Foundation (grant no. 6182021, Beijing, China), and the Earmarked Fund for Modern Agro-industry Technology Research System (grant no. CARS-36). The authors thank all the members of the Animal Molecular and Quantitative Genetics Laboratory in China Agricultural University (Beijing, China) for their assistance in the study design and data analysis. The authors have not stated any conflicts of interest.

REFERENCES


**ORCIDS**

Q. Z. Chen @ https://orcid.org/0000-0002-7737-476X
M. Y. Yang @ https://orcid.org/0000-0003-3872-4031
X. Q. Liu @ https://orcid.org/0000-0002-8160-015X
J. N. Zhang @ https://orcid.org/0000-0002-7967-8425
S. Y. Mi @ https://orcid.org/0000-0003-4950-680X
Y. J. Wang @ https://orcid.org/0000-0002-3605-8437
W. Xiao @ https://orcid.org/0000-0002-7909-8711
Y. Yu @ https://orcid.org/0000-0002-4524-0791


Chen et al.: GENES ASSOCIATED WITH SUPERNUMERARY TEATS

Journal of Dairy Science Vol. 105 No. 12, 2022