

Book Chapter

Genomic Identification, Evolution, and Expression Analysis of Bromodomain Genes Family in Buffalo

Qiang Fu*, Yangqing Lu, Ming Zhang and Junjun Zhang

State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Animal Reproduction Institute, Guangxi University, China

***Corresponding Author:** Qiang Fu, Email: gxfuq@gxu.edu.cn

Published **April 25, 2022**

This Book Chapter is a republication of an article published by Qiang Fu, et al. at Genes in January 2022. (Zhang, J.; Huang, L.; Zhang, P.; Huang, X.; Yang, W.; Liu, R.; Sun, Q.; Lu, Y.; Zhang, M.; Fu, Q. Genomic Identification, Evolution, and Expression Analysis of Bromodomain Genes Family in Buffalo. Genes 2022, 13, 103. <https://doi.org/10.3390/genes13010103>)

How to cite this book chapter: Qiang Fu, Yangqing Lu, Ming Zhang, Junjun Zhang. Genomic Identification, Evolution, and Expression Analysis of Bromodomain Genes Family in Buffalo. In: Fekadu Gadissa, editor. Prime Archives in Genetics: 2nd Edition. Hyderabad, India: Vide Leaf. 2022.

© The Author(s) 2022. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Author Contributions: writing—original draft preparation, J.Z.; data curation and formal analysis, L.H.; validation, P.Z.; methodology and software, X.H., W.Y., R.L., and Q.S.; project administration, Y.L.; project administration, M.Z.; writing—

review and editing, Q.F. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Natural Science Foundation of P.R. China (Grant No.31960660, 32160789).

Institutional Review Board Statement: All procedures involving animal treatment in this study were based on animal use guidelines of the Council for International Organizations of Medical Sciences and were approved by the Ethics Committee for Animal Experiments of Guangxi University.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Acknowledgments: We thank the smallholders for their assistance with the buffalo samples. We would like to thank Bullet Edits (<http://www.bulletedits.cn/>) for English language editing. We also would like to thank the anonymous reviewers for their constructive comments.

Conflicts of Interest: The authors declare no conflict of interest.

Abstract

Bromodomain (BRD) is an evolutionarily conserved protein–protein interaction module that is critical in gene regulation, cellular homeostasis, and epigenetics. This study aimed to conduct an identification, evolution, and expression analysis of the BRD gene family in the swamp buffalo (*Bubalus bubalis*). A total of 101 BRD protein sequences deduced from 22 BRD genes were found in the buffalo genome. The BRD proteins were classified into six groups based on phylogenetic relationships, conserved motifs, and conserved domains. The BRD genes were irregularly distributed in 13 chromosomes. Collinearity analysis revealed 20 BRD gene pairs that had remarkable homologous relationships between the buffalo and cattle, although no tandem or segmental duplication event was found in the buffalo BRD genes. Comparative transcriptomics

using a 10x sequencing platform analysis showed that 22 BRD genes were identified in the Sertoli cells (SCs) at different developmental stages of buffalo. Further, the mRNA expression levels of bromodomain and the extraterminal (BET) family in SCs at the pubertal stage were higher than that at the prepubertal stage of buffalo. However, the *SMARCA2*, *PHIP*, *BRD9*, and *TAF1* genes exhibited the opposite trend. The maturation process of SCs may be regulated by the BRD family members expressed differentially in SCs at different developmental stages of buffalo. In summary, our findings provide an understanding of the evolutionary, structural, and functional properties of the buffalo BRD family members, and further characterize the function of the BRD family in the maturation of SCs. It also provides a theoretical basis for further understanding in the future of the mechanism of SCs regulating spermatogenesis.

Keywords

Genomic Identification; BRD Family; Evolution; Expression Analysis; Buffalo

Introduction

Bromodomain (BRD) is a novel protein domain with approximately 110 amino acids [1,2], and was first discovered by John W. Tamkun and colleagues in the *Brahma/brm* gene of *Drosophila melanogaster* [3]. In the human genome, 61 different BRD genes encode 46 diverse BRD-containing proteins. The BRD-containing proteins belong to eight different families (BRD I–VIII) [4]. The BRD module consists of four α -helices connected by two divergent loop regions, named ZA and BC loops, that form a hydrophobic pocket, which contributes to the recognition of acetylated lysine [5]. These proteins act as a scaffold for transcriptional regulators, chromatin regulators, and chromatin-modifying enzymes [6]. Additionally, BRD proteins are expressed in various tissues, including the pancreas, testis, ovary, brain, lung, and kidney, as indicated by previous studies [7–10], and show extensive and complicated expression profiles. This suggests that BRD may have a context-dependent function in cellular homeostasis. Moreover, BRD-containing proteins can

regulate downstream gene expression, which makes them valuable targets for the study of fertility, neurological and immunological diseases, and cancer therapy [11–15].

Previous reports have shown that BRD-containing proteins have multiple physiological functions, including histone acetyltransferase activity, chromatin remodeling, transcriptional regulation, and co-activation. The most notable function of BRD-containing proteins is the transcriptional regulation of gene expression [16]. The BRD family proteins, such as bromodomain and the extraterminal (BET) family proteins, are crucial for cell development. The BRD testis-specific proteins (BRDT), BRD2, BRD3, and BRD4 are members of this family. Notably, BRDT is specifically expressed in the testis and is closely associated with spermatogenesis [17,18]. Moreover, members of the BET family are commonly expressed in Sertoli cells (SCs) [7]. Research has shown that inhibition of the BET family induces male contraception in mice [19]. Testicular SCs are the major somatic component of spermatogenic tubules and establish tight junctions to form a blood–testis barrier during puberty. SCs are important for maintaining testicular homeostasis during spermatogenesis [20,21]. However, the BRD family has not been extensively studied in the context of testicular SCs.

The swamp buffalo is an adaptable livestock with many outstanding traits and is considered to be one of the most potentially valuable domestic animals by the Food and Agriculture Organization of the United Nations. Recent advances have further revealed the complex and versatile functions of BRD proteins in regulating gene expression and mediating protein–protein interactions [10,22]. However, there are only a few studies on the BRD gene family in buffalo, which has hindered the exploitation of the potential value of buffalo. Thus, it is crucial to reveal whether the BRD gene family affects growth and development, transcriptional regulation, epigenetics, and disease treatment in buffalo. Further, the effect of genetic differences of the BRD family in buffalo SCs maturation and function during spermatogenesis remains unclear. Hence, the

expression profile of the BRD family in buffalo SCs needs to be elucidated.

Bioinformatics, combined with next-generation sequencing, can provide high-throughput genomic data. These data can be used to identify the genes for the favorable traits in buffalo, for instance, the systematic analysis of candidate genes for buffalo milk production and fertility traits [23,24], whole-genome sequencing and characterization for selective breeding in buffalo [25], and characterization of the diacylglycerol acyltransferases gene family, collagen gene family, and heat-shock protein gene family [26–28]. The above studies, together with a nearly complete genome map of the buffalo [29–32], make it possible to study gene families at the genome-wide level in buffalo.

Based on present research progress, we applied an integrated bioinformatics approach to identify the BRD gene family and analyze its evolutionary relationships, sequence characteristics, chromosomal locations, collinearity, and expression levels. We believe that the findings of this study will provide a fundamental basis for completing the buffalo genetic map. Furthermore, it will help in elucidating the roles of all BRD family members in the gene regulation of SCs maturation, even spermatogenesis.

Materials and Methods

Identification of the BRD Genes in Buffalo

Genome-wide data of seven mammals, including human ([GRCh38.p13](#)), mouse ([GRCm39](#)), rat ([Rnor_6.0](#)), cattle ([ARS-UCD1.2](#)), buffalo ([UOA_WB_1](#)), pig ([Sscrofa11.1](#)), and goat ([ARS1](#)) were retrieved from the National Center for Biotechnology Information (NCBI) website (<https://www.ncbi.nlm.nih.gov/>, accessed on 27 October 2020) [33] to identify the BRD family of the buffalo. The Hidden Markov Model (HMM) archive of bromodomain ([PF00439](#)) was downloaded from the Pfam website (<http://pfam.xfam.org/>, accessed on 27 October 2020) [34]. The HMMER program [35–37] was used to search the buffalo dataset to find proteins containing the BRD domain (E-value, 1×10^{-5}). In addition, we employed the ClustalW program to align multiple sequences and

independently construct the buffalo HMM profile. For further validation, we compared the BRD protein of buffalo with that of other mammals on the NCBI database using BLAST. The BRD protein sequences of seven species were used to construct a neighbor-joining (NJ) phylogenetic tree using the MEGA X software [38].

Sequence Analysis of BRD Family

To characterize the physicochemical properties of BRD proteins, the ExPASy server (<https://www.expasy.org/>, accessed on 16 September 2021) [39] was used to predict the molecular weight (MW), isoelectric point (*pI*), grand average of hydrophilicity (GRAVY), and the amino acid sequences of the buffalo BRD protein family [40,41]. The BUSCA online program (<http://busca.biocomp.unibo.it/>, accessed on 26 September 2021) [42] was applied to predict the subcellular location of the BRD family. Gene structures were measured and visualized using Tbttools (v1.09854) [43]. Exon and intron structure analysis was performed by the Gene Structure Display Server tool [44]. The conserved motifs of the buffalo BRD proteins were analyzed by the MEME program (<http://meme.nbcr.net/>, accessed on 10 November 2020) [45], with a maximum number of motifs set at 10 [46–48]. Conserved domains of protein were obtained by the Conserved Domains tool provided by NCBI (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi/>, accessed on 10 November 2020) [49].

Chromosomal locations of BRD genes were obtained from their genomic dataset. Visualization of replication events and collinearity of BRD genes were implemented using the MCScanX [50,51]. First, BLAST searches were performed on 58,204 buffalo protein sequences (E-value, 1×10^{-5}). Then, the information on corresponding chromosome length and BRD gene locations was retrieved from the genomic sequence files of buffalo and cattle and was submitted to the MCScanX program. The chromosomal distribution and the collinearity of BRD genes between buffalo and cattle were visualized by Tbttools.

Comparative Transcriptomic Analysis of SCs

The single-cell transcriptomes of SCs from prepubertal (3 months) and pubertal (24 months) buffalo were analyzed ([GSE190477](#)). Briefly, the expression matrix was generated using the count module of the Cell Ranger analysis pipeline, and subsequent analysis was performed using the R package “Seurat [50,51].

Isolation of SCs from Buffalo Testicular Material

The isolation protocol of SCs was modified according to previous studies [54–58]. Buffalo testes were obtained from a local slaughterhouse in Nanning and were transported to the laboratory in saline maintained at 37 °C. Each group had three biological replicates. Testicular tissues were peeled off and cut into an erosive mass on a super clean bench, while the testicular tissues were washed three times with PBS containing penicillin and streptomycin (Solarbio, Beijing, China). We used a two-step enzymatic digestion method to isolate SCs from buffalo testicular material. First, spermatogenic tubules were obtained by digesting the testicular material with DMEM/F12 solution containing collagenase type IV (0.5 mg/ml; Worthington Biochemical Corporation, NJ, USA) and Dnase I (25 µg/ml; Sigma Life Science, MO, USA) at 37 °C for approximately 12 min. Then, a second enzymatic digestion was performed using DMEM/F12 solution containing trypsin (0.25%; Gibco, MA, USA) and Dnase I (25 µg/ml), and incubated for approximately 10 min at 37 °C. After filtering through a cell strainer, the mixture of germ cells and SCs were cultured in a DMEM/F12 medium containing 10% FBS for 4 h (37 °C, 5% CO₂). Finally, the unadhered germ cells were removed. The remaining adhered cells were the SCs.

Real-Time Quantitative PCR

The total RNA from buffalo SCs was extracted using the RNAiso Plus kit (Takara, Beijing, China). The quality of the total RNA was evaluated by a microplate spectrophotometer (Epoch, BioTek, VT, USA). Reverse transcription of the total RNA (≈ 1.5

μg) to cDNA was done using a reverse transcriptase kit (Vazyme, Nanjing, China). Real-time quantitative PCR (qRT-PCR) analysis was performed to amplify the target genes using SYBR qPCR Master Mix (Vazyme, Nanjing, China), and the reaction was conducted on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad, CA, USA). Reactions were performed in a final volume of 20 μL with 0.4 μL of each primer (≈4 μM), 2 μL of cDNA (≈150 ng), 10 μL of SYBR qPCR Master Mix, and 7.2 of μL ddH₂O. The reaction was run with the following parameters: 95 °C, 30 s, followed by 95 °C, 10 s and 60 °C, 30 s for 40 cycles. Melting curves were acquired using the default parameters of the CFX96 Touch Real-Time PCR Detection System. Negative qRT-PCR controls were executed by ddH₂O and RNA, instead of cDNA. Every reaction was repeated three times and was normalized against GAPDH expression [59–62]. The $2^{-\Delta\Delta CT}$ approach [63] was used to calculate the relative expression levels of the chosen BRD genes in buffalo. Table S1 lists the primers used in the qRT-PCR experiments.

Statistical Analysis

Statistical analyses were conducted using Prism 9 software (GraphPad Software, La Jolla, CA, USA). The values are expressed as mean ± S.D. Grouped differences were assessed by two-way analysis of variance (ANOVA), followed by the Holm–Šidák test for multiple comparisons. Differences were considered significant for *p* values < 0.05.

Results

Identification of BRD Genes and Proteins

Twenty-two BRD genes were identified from the entire buffalo genome using BLAST and HMMER programs. The deduced protein products, including 101 non-redundant protein sequences, were encoded by the 22 BRD genes. The relevant information is in Table S2. Analysis of physicochemical properties indicated that the open reading frames of the isoforms of different BRD genes ranged from 435 bp to 9591 bp, encoding proteins of 144 to 3196 amino acid residues with the

predicted MW of 16.13–353.47 kDa. The *pI* of these protein isoforms ranged from 4.49 to 9.40, while the GRAVY range was between –1.216 and –0.312. In addition, phylogenetic analysis revealed that the representative BRD proteins of the seven mammalian species could be divided into six branches (Figure 1). The genes displayed in the sixth group included more nodes, but were grouped for a better presentation of the results. The maximum ($N = 12$) and minimum ($N = 2$) number of BRD proteins were enriched in the sixth and fifth groups, respectively. The phylogenetic analysis indicated that the buffalo BRD protein family was closely connected to the other six mammalian species. With respect to genetic evolution, the buffalo was the closest to cattle but the farthest from rats and mice.

Sequence Analysis of Buffalo BRD Family

The motif pattern, gene structure, and conserved domain were investigated based on the genetic relationships of the buffalo BRD family (Figure 2 and 3). Figure 2B shows that BRD proteins include 10 conserved motifs. A Pfam search revealed that motifs 1–4 were annotated as BRD modules with 28–50 amino acids (Table 1). This result was supported by the identified conserved domains of BRD proteins, according to the Conserved Domain Database of NCBI (Figure 2C). Interestingly, BET and BRD4 domains were also detected in some BRD proteins with the precise characteristics of the BET family, suggesting that the BET family proteins containing the double BRD modules might have a superior function. In the same group, gene structure analysis showed that buffalo BRD genes had similar numbers of exons and introns, despite the structures of the coding sequence (CDS) and untranslated region (UTR) being different (Figure 3). This result indicates that different clades have distinct exon patterns, confirming our classification. Furthermore, members of the BRD family with a more distant genetic relationship have a higher structural and functional diversity.

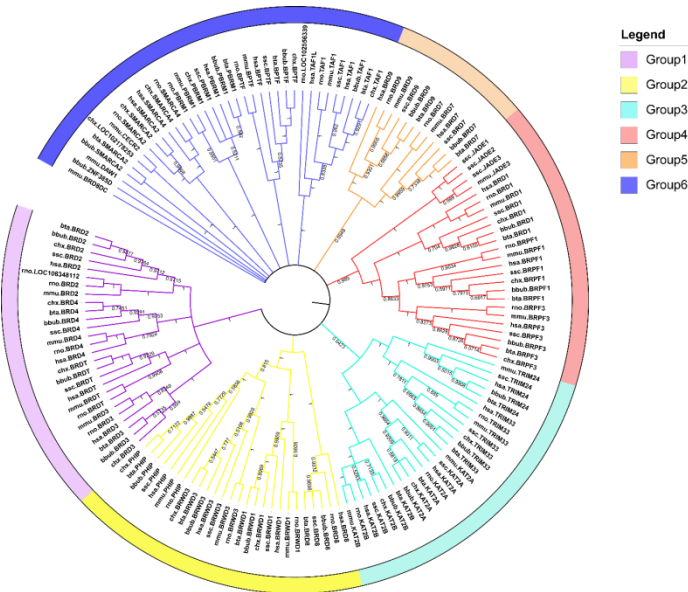


Figure 1: NJ tree of BRD proteins in seven mammals. Different-colored clades and strips indicate different groups. Note: buffalo: bbub; mouse: mmu; rat: rno; human: hsa; goat: chx; cow: bta; pig: ssc.

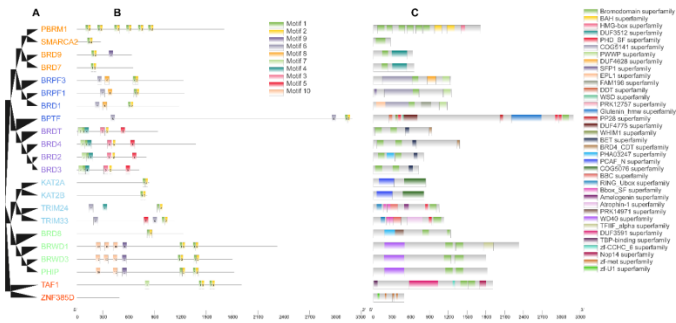


Figure 2: Evolutionary relationships, motif pattern, and conserved domains of BRD protein family in buffalo. (A) Phylogenetic tree of 22 BRD proteins. (B) Motif pattern of BRD proteins. Ten presumptive motifs are indicated with boxes marked different colors. Refer to Table 1 for more information of the motifs. (C) Distributions of conserved domains in BRD proteins.



Figure 3: The UTR/CDS structure of BRD genes in buffalo. The green box, black line, and yellow box represent UTR, intron, and CDS, respectively.

Table 1: Ten distinct motifs generally detected in buffalo BRD family.

Motif	Protein Sequence	Length	Pfam Domain
MEME-1	APDYYKIIKKPMDLSTIKE RLENNYYQS	28	BRD
MEME-2	FVADVRLIFSNCRKYNPP DSEVYKAAKKL	29	BRD
MEME-3	QLKHCSGILKEMLSKKHA AYAWPFYKPV DVEALGL HDYHDIKHPMDLST	50	BRD
MEME-4	ASECIQDFNTMFTNCYIY NKP GDDIVLMAQALEKJF LQKVAQMPQEE	47	BRD
MEME-5	PMSYDEKRQLSLDINKLP GEKLG R VVHIIQSREPSLR DSNPDEIEIDFET	50	BET
MEME-6	DAVCCVCLDGECQNSNVI LFCDMCNLAVHQECYGV PYIPEGQWLCRRCLQ	50	---
MEME-7	NPPPPEVSNPKKPGRLTN QLQYLQKVVLKALWKHQ FAWPFQQPVDVAVKLN	50	---
MEME-8	VCFANTVFLEPIDGIDNIPP ARWKLTCYICKQKGVGA CIQCHKANCYTAF	50	---
MEME-9	HFACTDSHGHLIFGFGCS KPYEKIPDQMFFHTDYRP LIRDANNYVLDEQ	50	---
MEME-10	RGHSAEISDMAVNYENT MIAAGSCDKIIRVWCLRT CAPVAVLQGHASIT	50	---

Chromosomal Distribution and Collinearity Analysis of BRD Genes

As shown in Figure 4, chromosomal distribution analysis showed that all identified buffalo BRD genes were randomly distributed on 13 chromosomes (Figure 4A), while cattle BRD genes were randomly distributed on 14 chromosomes (Figure 4B). The chromosomal distribution of BRD genes was not similar between buffalo and cattle, except for *BRWD1* and *KAT2B* on chromosome 1 and *BRD7* on chromosome 18. Notably, the *TAF1* and *BRWD3* of the two species were distributed at similar positions on the X chromosome, suggesting that genes on the sex chromosome are highly conserved.

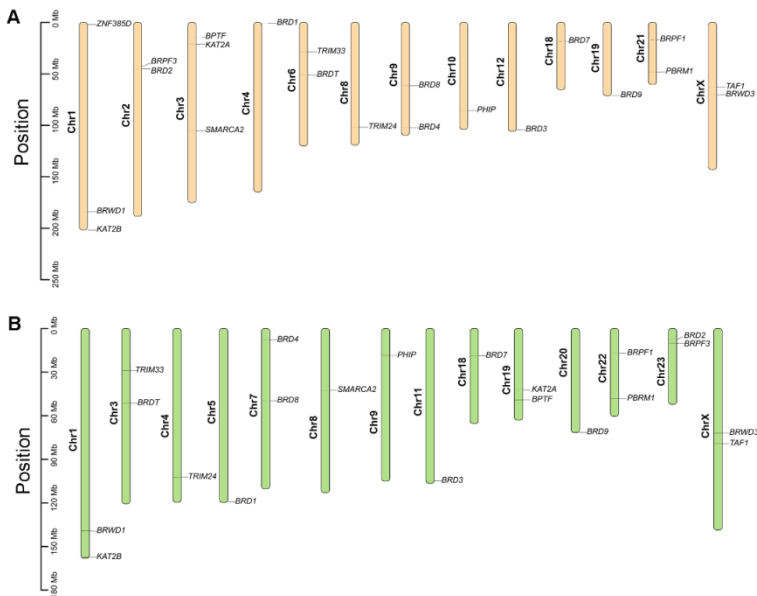


Figure 4: Chromosomal distribution of BRD genes in buffalo and cattle. (A) Position of BRD gene on buffalo chromosome. (B) Position of BRD gene on cattle chromosome. The ellipses, which are marked with orange and green colors, represent buffalo and cattle chromosomes, respectively.

To study the genetic evolutionary process of the BRD family, we investigated the duplication events in the BRD gene families of buffalo and cattle (Figure 5). Tandem and segmental

duplication events were not present in buffalo BRD genes (Figure 5A), which means that the BRD gene in buffalo is relatively conserved. Collinearity analysis identified a total of 34,862 collinear genes in the buffalo and cattle genomes, covering 82.31% of the total number of genes. Interestingly, despite having a different number of chromosomes, a great number of chromosomal homologies were found between buffalo and cattle ($2n = 50$ and 60 , respectively, where n is the haploid number of chromosomes). As shown in Figure 5B, 20 pairs of BRD genes were homologous between the two species, detailed information for which is given in Table S3. The results suggest that the 20 pairs of homologous BRD genes had the same ancestral chromosome and perhaps belonged to orthologous genes. However, some differences in the gene order occurred between buffalo and cattle over evolution.

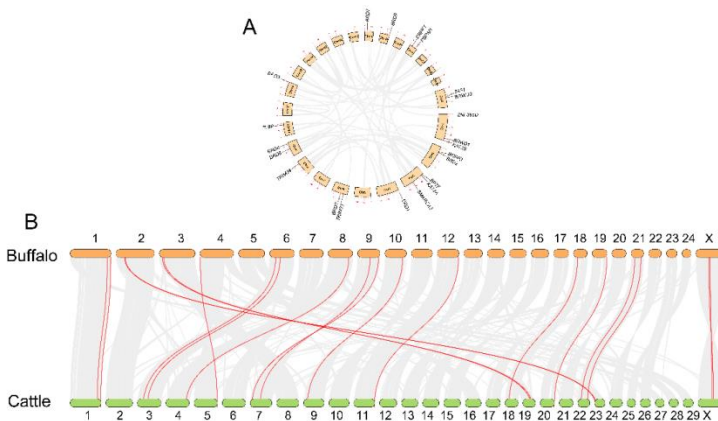


Figure 5: Gene duplication of buffalo genome (A) and collinear analysis of buffalo and cattle genome (B). Collinear genes are linked by gray lines, and collinear BRD genes are linked by red lines.

Comparative Transcriptional Analysis of BRD Genes in SCs

Using the single-cell RNA-sequencing data, we analyzed the expression differences of BRD genes in buffalo SCs at different developmental stages. SCs of prepubertal and pubertal buffalo were defined as immature and mature cells, respectively [64,65].

The expression levels of most BRD genes were significantly different between immature and mature SCs. *BRDT* and other BET family members (*BRD2*, *BRD3*, and *BRD4*) had a significantly higher expression in the mature group over the immature group. Nevertheless, in immature SCs, the expression of *SMARCA2*, *PHIP*, *BRD9*, and *TAF1* was markedly higher than in mature SCs. (Figure 6A, B). The distribution of immature and mature SCs and the features of BRD genes in immature and mature SCs are shown in Figures S1 and S2. We detected 18 upregulated and 4 downregulated BRD genes, which accounted for 3.04% and 0.55% of the identified upregulated and downregulated genes in mature SCs compared with immature SCs, respectively (Figure S3). The morphological characteristics and identification of immature and mature SCs are displayed in Figures S4 and S5, respectively. Further, qRT-PCR analysis showed that the mRNA expression trends of the chosen *SMARCA2*, *PHIP*, *BRD9*, *TAF1*, and BET family genes in buffalo were consistent with the RNA-Sequencing data (Figure 6C), suggesting that these genes have specific spatiotemporal expression patterns and are critical in the maturation of SCs.

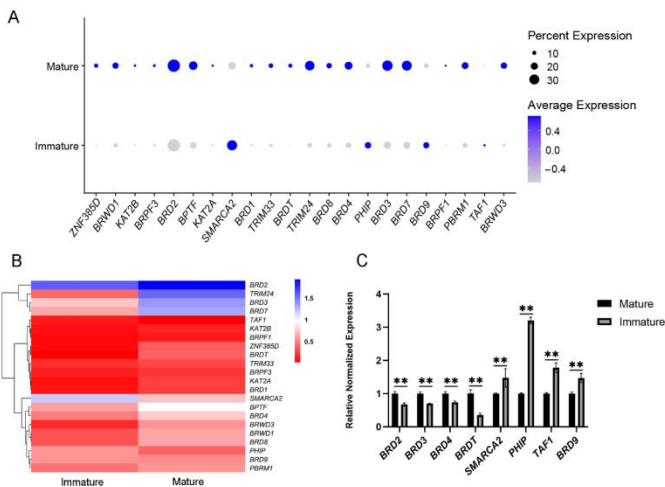


Figure 6: Dot plot (A) and heatmap (B) of BRD genes in immature and mature SCs, and verification of chosen buffalo BRD genes by qRT-PCR (C). The dot size represents the numbers of genes, and the color shade represents the expression level of genes in SCs. * Significant difference. ** Significant difference.

Discussion

Physicochemical Properties and Phylogenetic Analysis of BRD Families

BRD, an evolutionarily conserved protein–protein interaction module, plays a vital role in gene regulation and maintenance of cellular homeostasis [6]. BRD-containing proteins are critical as elements of transcription factor complexes and determinants of epigenetic memory [66]. Currently, functional studies on the buffalo BRD gene family are limited. In this study, we identified 101 BRD protein sequences based on the complete buffalo genome map using bioinformatics analysis. The identified BRD protein sequences corresponded to the 22 BRD genes in buffalo.

Characterization of the physicochemical properties of proteins encoded by different gene families is essential to elucidate their functions and characteristics. The predicted *pI* indicates that the identified buffalo BRD proteins are a mix of acidic and alkaline proteins, which implies the functional diversity among them. GRAVY values were used to predict the water and protein interaction. A positive GRAVY score suggests a globally hydrophobic protein, while a negative one suggests solubility. In our study, all members of the BRD protein family had negative GRAVY scores, indicating that these proteins are soluble. These physicochemical properties reflect the functional diversity of BRD isoforms. These results further provide a reference for a better understanding of the function of the BRD protein family.

The evolutionary relationships between the BRD proteins of buffalo with six other representative mammals were investigated. We classified the evolutionary relationships into six groups based on the structural topology and sequence similarity. The classification of the human BRD protein family by Filippakopoulos and Fujisawa was used [4,6]. It is noteworthy that, although the identification of the BRD family in cattle failed to categorize *BRDT*, the distribution of cattle *BRDT* on chromosome 3 was still demonstrated in the results. This might be because of the strict E-value threshold in our identification method. Further, *ZNF385D* was excluded upon gene structural analysis because no conserved motif was

detected in it, which reflects that those genes and proteins have specific sequence characteristics among different species. It also indirectly illustrates that the present identification method is not perfect, and a more scientific analysis method should be established.

Sequence and Structure Analysis of BRD Family

Conservative domain analysis showed that all identified BRD proteins, except ZNF385D, contained the BRD domain at least, which is supported by previous studies [4]. Interestingly, the characteristic BRD domains, such as BET and BRD4_CDT [11,66], were also detected in the BRD proteins. In addition, it is also worth noting that BAH and PHD_SF [6,67] were capable of mediating protein interactions as epigenetic readers. From the gene structure analysis, we found significant differences in the distributions of introns and exons, even though these genes have close evolutionary relationships, but the major coding sequences of genes were similar. This is contrary to the belief that the difference in the length and distribution of CDS and UTR should be the primary reason for the gene structural disparity. This result corroborates studies of other gene families [26,27]. Overall, the diversity in the structural arrangement of nucleic acid and amino acid sequences within the BRD family implies that they have different functions and may play non-redundant roles in buffalo.

Chromosomal Distribution and Collinearity Analysis of BRD Genes

Chromosome doubling, chromosome segment duplication, tandem duplication, and gene translocation are four types of gene duplication events that can lead to gene complexity and novel gene functions during genetic evolution [68–70]. The chromosomal distribution of BRD genes showed that 22 genes were located on 13 chromosomes in buffalo, while 21 genes were located on 14 different chromosomes in cattle. The 22 BRD genes were unevenly distributed across the 13 chromosomes in buffalo. *TAF1* and *BRWD3* are distributed on sex chromosomes, and the specificity of their distribution on the

X chromosome may be related to sex-linked inheritance. Notably, we detected homologous *BRDT* in cattle and buffalo in the collinearity analysis results. We still showed the chromosomal localization of *BRDT* in cattle, although our identification method failed to find it. This result is probably due to the discrepancy of the selected reference genome. Further, we could not detect tandem and segmental duplication, but the relevant results were still provided for access (Tables S4 and S5). Tandem and segment duplication mainly contribute to the acceleration of gene family expansion and genomic complexity evolution [71–73]. This result may indicate that the buffalo BRD family genes are relatively conserved during the evolutionary process.

Based on the results of the collinearity analysis, we observed that most of the chromosomes between the buffalo and cattle are homologous. Further, segmental duplication events occurred during the amplification of BRD genes in the buffalo and cattle genomes, which may belong to orthologous genes. A similar phenomenon was found in the genome map of the buffalo [30,31,74]. The collinear chromosomes of the two species descended from a common ancestral chromosome. Except for the *BRWD1* and *KAT2B* genes located on chromosome 1, *BRWD3* and *TAF1* on the X chromosome, and *BRD7* on chromosome 18, most BRD genes were located in inconsistent chromosome locations. Each pair of BRD genes between buffalo and cattle was either syntenic (located on the same chromosome) or collinear (conserved on different chromosomes). Over many years of evolution, differences in gene order may have occurred due to chromosomal rearrangements. Overall, the considerable chromosomal homology between buffalo and cattle raises the potential for researching genetic innovation and structural variation in the buffalo BRD family.

Expression Analysis of BRD Genes in Immature and Mature SCs

We found a total of 22 BRD genes at the different developmental stages in buffalo. Notably, the testicular BRD gene family showed significantly different expression levels at

different developmental stages in the same species, suggesting that these genes have specific spatiotemporal expression patterns. This is consistent with the findings of previous studies [7]. During early developmental processes, SCs begin to form, proliferate, and differentiate, which marks the transition from the testicular cord to the spermatogenic tubules. However, at the beginning of puberty, SCs stop proliferating, and undergo novel changes in morphology and function [75]. In particular, spermatogenic capacity increases as the numbers of SCs increase [76,77]. Hence, it is crucial to study the transformation processes of immature to mature SCs.

BRD2 can function as a transcriptional coactivator or corepressor protein [78,79]. BRD3 interacts with the acetylated transcription factor GATA1 and is important in erythropoiesis by regulating erythrocyte target genes [80]. BRD4 is an all-basic requirement BET protein in mammalian cells [81]. BRDT is a critical factor in transcriptional prolongation and plays a key role in spermatogenesis [14,19].

In our study, the expression of *BRDT* was significantly higher in the mature group than in the immature group of SCs. Interestingly, other BET family members (*BRD2*, *BRD3*, and *BRD4*) also showed the same expression trend. This indicates that they may have specific roles at different developmental stages, perhaps contributing to the transformation of immature SCs into mature SCs, which interact with germ cells to support spermatogenesis. However, in the immature SCs, the expression of *SMARCA2*, *PHIP*, *BRD9*, and *TAF1* was significantly higher than in the mature group. *SMARCA2* and *BRD9* are components of the SWI/SNF complex, which contributes to the coordinated regulation of gene expression programs [82,83], while *PHIP* and *TAF1* are transcriptional coregulators that control gene expression to regulate a variety of biological processes, such as signal transduction, cell proliferation, cell cycle, and apoptosis [84–87]. High expression of these genes may contribute towards the maintenance of certain biological characteristics of immature SCs. The specific expression patterns of BRD family genes in the above results demonstrate that their expression is strictly regulated during buffalo growth, and these findings are validated

by qRT-PCR. Hence, our findings suggest that these BRD genes exhibit distinct biological functions at different developmental stages, which will help to dissect the potential role of these genes in the maturation of SCs. More importantly, it provides a theoretical basis for studying the contribution of the BRD family to the maintenance of testicular homeostasis and the regulation of spermatogenesis.

Conclusions

In this study, we completely annotated the BRD family in buffalo by bioinformatics. The differences in physicochemical properties and structures of the identified 22 BRD members may be related to the diversity of functions. Collinearity analysis indicated that the buffalo BRD family is relatively conserved during evolution. In addition, the comparative transcriptional analysis showed that BRD genes have different expression patterns in testicular SCs at different developmental stages, which may be correlated with the maturation of SCs and the precise regulation necessary for spermatogenesis. Hence, this study provides valuable information on the buffalo BRD gene family and will contribute towards understanding the role of BRD family members in affecting the maturation of SCs and regulating spermatogenesis.

References

1. Haynes SR, Dollard C, Winston F, Beck S, Trowsdale J, et al. The bromodomain: A conserved sequence found in human, *Drosophila* and yeast proteins. *Nucleic Acids Res.* 1992; 20: 2603.
2. Jeanmougin F, Wurtz JM, Le Douarin B, Chambon P, Losson R. The bromodomain revisited. *Trends Biochem Sci.* 1997; 22: 151–153.
3. Tamkun JW, Deuring R, Scott MP, Kissinger M, Pattatucci AM, et al. *brhma*: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell.* 1992; 68: 561–572.
4. Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JP, et al. Histone recognition and large-scale structural

- analysis of the human bromodomain family. *Cell*. 2012; 149: 214–231.
5. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, et al. Structure and ligand of a histone acetyltransferase bromodomain. *Nature*. 1999; 399: 491–496.
 6. Fujisawa T, Filippakopoulos P. Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. *Nat. Rev. Mol. Cell Biol*. 2017; 18: 246–262.
 7. Shang E, Salazar G, Crowley TE, Wang X, Lopez RA, et al. Identification of unique, differentiation stage-specific patterns of expression of the bromodomain-containing genes Brd2, Brd3, Brd4, and Brdt in the mouse testis. *Gene Expr. Patterns*. 2004; 4: 513–519.
 8. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, et al. A draft map of the human proteome. *Nature*. 2014; 509: 575–581.
 9. Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, et al. Mass-spectrometry-based draft of the human proteome. *Nature*. 2014; 509: 582–587.
 10. Wang N, Wu R, Tang D, Kang R. The BET family in immunity and disease. *Signal Transduct. Target. Ther*. 2021; 6: 23.
 11. Muller S, Filippakopoulos P, Knapp S. Bromodomains as therapeutic targets. *Expert Rev. Mol. Med*. 2011; 13: e29–e29.
 12. Belkina AC, Denis G. V BET domain co-regulators in obesity, inflammation and cancer. *Nat. Rev. Cancer*. 2012; 12: 465–477.
 13. Wang CY, Filippakopoulos P. Beating the odds: BETs in disease. *Trends Biochem Sci*. 2015; 40: 468–479.
 14. Gaucher J, Boussouar F, Montellier E, Curtet S, Buchou T, et al. Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J*. 2012; 31: 3809–3820.
 15. Chavez DR, Lee PC, Comizzoli P. Oocyte Meiotic Competence in the Domestic Cat Model: Novel Roles for Nuclear Proteins BRD2 and NPM1. *Front. cell Dev. Biol*. 2021; 9: 670021.
 16. Florence B, Faller D. V You bet-cha: a novel family of transcriptional regulators. *Front Biosci*. 2001; 6: D1008-18.
 17. Bryant JM, Berger SL. Low-hanging fruit: targeting Brdt in

- the testes. *EMBO J.* 2012; 31: 3788–3789.
18. Berkovits BD, Wolgemuth DJ. The role of the double bromodomain-containing BET genes during mammalian spermatogenesis. *Curr Top Dev Biol.* 2013; 102: 293–326.
 19. Matzuk MM, McKeown MR, Filippakopoulos P, Li Q, Ma L, et al. Small-molecule inhibition of BRDT for male contraception. *Cell.* 2012; 150: 673–684.
 20. França LR, Hess RA, Dufour JM, Hofmann MC, Griswold MD. The Sertoli cell: one hundred fifty years of beauty and plasticity. *Andrology.* 2016; 4: 189–212.
 21. Ni FD, Hao SL, Yang WX. Multiple signaling pathways in Sertoli cells: recent findings in spermatogenesis. *Cell Death Dis.* 2019; 10: 541.
 22. Cheung KL, Kim C, Zhou MM. The Functions of BET Proteins in Gene Transcription of Biology and Diseases. *Front. Mol. Biosci.* 2021; 8: 728777.
 23. Du C, Deng T, Zhou Y, Ye T, Zhou Z, et al. Systematic analyses for candidate genes of milk production traits in water buffalo (*Bubalus Bubalis*). *Anim Genet.* 2019; 50: 207–216.
 24. Vohra V, Chhotaray S, Gowane G, Alex R, Mukherjee A, et al. Genome-Wide Association Studies in Indian Buffalo Revealed Genomic Regions for Lactation and Fertility. *Front. Genet.* 2021; 12: 696109.
 25. Rehman SU, Hassan F, Luo X, Li Z, Liu Q. Whole-Genome Sequencing and Characterization of Buffalo Genetic Resources: Recent Advances and Future Challenges. *Animals.* 2021; 11.
 26. Liu J, Wang Z, Li J, Li H, Yang L. Genome-wide identification of Diacylglycerol Acyltransferases (DGAT) family genes influencing Milk production in Buffalo. *BMC Genet.* 2020; 21: 26.
 27. Lu X, Duan A, Liang S, Ma X, Deng T. Genomic Identification, Evolution, and Expression Analysis of Collagen Genes Family in Water Buffalo during Lactation. *Genes (Basel).* 2020; 11: 515.
 28. Rehman SU, Nadeem A, Javed M, Hassan FU, Luo X, et al. Genomic Identification, Evolution and Sequence Analysis of the Heat-Shock Protein Gene Family in Buffalo. *Genes (Basel).* 2020; 11: 1388.

29. Low WY, Tearle R, Bickhart DM, Rosen BD, Kingan SB, et al. Chromosome-level assembly of the water buffalo genome surpasses human and goat genomes in sequence contiguity. *Nat. Commun.* 2019; 10: 260.
30. Williams JL, Iamartino D, Pruitt KD, Sonstegard T, Smith TPL, et al. Genome assembly and transcriptome resource for river buffalo, *Bubalus bubalis* (2n = 50). *Gigascience.* 2017; 6: 1–6.
31. Mintoo AA, Zhang H, Chen C, Moniruzzaman M, Deng T, et al. Draft genome of the river water buffalo. *Ecol. Evol.* 2019; 9: 3378–3388.
32. Deng T, Pang C, Lu X, Zhu P, Duan A, et al. De Novo Transcriptome Assembly of the Chinese Swamp Buffalo by RNA Sequencing and SSR Marker Discovery. *PLoS One.* 2016; 11: e0147132–e0147132.
33. National Center for Biotechnology Information. Available online at: <https://www.ncbi.nlm.nih.gov>.
34. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* 2019; 47: D427–D432.
35. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, et al. HMMER web server: 2018 update. *Nucleic Acids Res.* 2018; 46: W200–W204.
36. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 2011; 39: W29–W37.
37. Prakash A, Jeffryes M, Bateman A, Finn RD. The HMMER Web Server for Protein Sequence Similarity Search. *Curr Protoc Bioinforma.* 2017; 60L 3 15 1-3 15 23.
38. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 2018; 35: 1547–1549.
39. Bjellqvist B, Basse B, Olsen E, Celis JE. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis.* 1994; 15: 529–539.
40. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, et al. ExPASy: The proteomics server for in-depth protein

- knowledge and analysis. *Nucleic Acids Res.* 2003; 31: 3784–3788.
41. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, et al. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 2012; 40: W597–W603.
 42. Savojardo C, Martelli PL, Fariselli P, Profiti G, Casadio R. BUSCA: an integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res.* 2018; 46: W459–W466.
 43. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant.* 2020; 13: 1194–1202.
 44. Hu B, Jin J, Guo AY, Zhang H, Luo J, et al. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics.* 2015; 31: 1296–1297.
 45. Meme tests and motifs, many other useful info from NCBR.net. Available online at: <http://meme.nbc.net>.
 46. Bailey TL, Gribskov M. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics.* 1998; 14: 48–54.
 47. Bailey TL, Williams N, Misleh C, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 2006; 34: W369–W373.
 48. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 2009; 37: W202–W208.
 49. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 2020; 48: D265–D268.
 50. Wang Y, Tang H, Debarry JD, Tan X, Li J, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012; 40: e49–e49.
 51. Wang Y, Li J, Paterson AH. MCScanX-transposed: detecting transposed gene duplications based on multiple colinearity scans. *Bioinformatics.* 2013; 29: 1458–1460.
 52. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 2017; 8: 14049.

53. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 2018; 36: 411–420.
54. Zhang P, He W, Huang Y, Xiao K, Tang Y, et al. Proteomic and phosphoproteomic profiles of Sertoli cells in buffalo. *Theriogenology.* 2021; 170: 1–14.
55. Kala S, Kaushik R, Singh KP, Kadam PH, Singh MK, et al. In vitro culture and morphological characterization of prepubertal buffalo (*Bubalus bubalis*) putative spermatogonial stem cell. *J. Assist. Reprod. Genet.* 2012; 29: 1335–1342.
56. Pramod RK, Mitra A. In vitro culture and characterization of spermatogonial stem cells on Sertoli cell feeder layer in goat (*Capra hircus*). *J. Assist. Reprod. Genet.* 2014; 31: 993–1001.
57. Li TT, Geng SS, Xu HY, Luo AL, Zhao PW, et al. Effects of different culture systems on the culture of prepubertal buffalo (*Bubalus bubalis*) spermatogonial stem cell-like cells in vitro. *J. Vet. Sci.* 2020; 21: e13–e13.
58. Zhang H, Liu B, Qiu Y, Fan J feng, Yu S. jiu Pure cultures and characterization of yak Sertoli cells. *Tissue Cell.* 2013; 45: 414–420.
59. Barber RD, Harmer DW, Coleman RA, Clark BJ. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol. Genomics.* 2005; 21: 389–395.
60. Li Q, Domig KJ, Ertle T, Windisch W, Mair C, et al. Evaluation of potential reference genes for relative quantification by RT-qPCR in different porcine tissues derived from feeding studies. *Int. J. Mol. Sci.* 2011; 12: 1727–1734.
61. Klein C, Rutllant J, Troedsson MH. Expression stability of putative reference genes in equine endometrial, testicular, and conceptus tissues. *BMC Res. Notes.* 2011; 4: 120.
62. Gong ZK, Wang SJ, Huang YQ, Zhao RQ, Zhu QF, et al. Identification and validation of suitable reference genes for RT-qPCR analysis in mouse testis development. *Mol. Genet. Genomics.* 2014; 289: 1157–1169.
63. Livak KJ, Schmittgen TD. Analysis of Relative Gene

- Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods*. 2001; 25: 402–408.
64. Curtis SK, Amann RP. Testicular Development and Establishment of Spermatogenesis in Holstein Bulls. *J. Anim. Sci.* 1981; 53: 1645–1657.
 65. Rawlings N, Evans ACO, Chandolia RK, Bagu ET. Sexual Maturation in the Bull. *Reprod. Domest. Anim.* 2008; 43: 295–301.
 66. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, et al. Selective inhibition of BET bromodomains. *Nature*. 2010; 468: 1067–1073.
 67. Nakamura T, Blechman J, Tada S, Rozovskaia T, Itoyama T, et al. huASH1 protein, a putative transcription factor encoded by a human homologue of the *Drosophila ash1* gene, localizes to both nuclei and cell-cell tight junctions. *Proc. Natl. Acad. Sci. U. S. A.* 2000; 97: 7284–7289.
 68. Freeling M. Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome, Segmental, or by Transposition. *Annu. Rev. Plant Biol.* 2009; 60: 433–453.
 69. Conant GC, Wolfe KH. Turning a hobby into a job: How duplicated genes find new functions. *Nat. Rev. Genet.* 2008; 9: 938–950.
 70. He X, Zhang J. Gene Complexity and Gene Duplicability. *Curr. Biol.* 2005; 15: 1016–1021.
 71. Liu GE, Ventura M, Cellamare A, Chen L, Cheng Z, et al. Analysis of recent segmental duplications in the bovine genome. *BMC Genomics*. 2009; 10: 571.
 72. Cheung J, Wilson MD, Zhang J, Khaja R, MacDonald JR, et al. Recent segmental and gene duplications in the mouse genome. *Genome Biol.* 2003; 4: R47–R47.
 73. Dennis MY, Eichler EE. Human adaptation and evolution by segmental duplication. *Curr. Opin. Genet. Dev.* 2016; 41: 44–52.
 74. Bovine Genome S, Analysis C, Elsik CG, Tellam RL, Worley KC, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 2009; 324: 522–528.
 75. Sharpe RM, McKinnell C, Kivlin C, Fisher JS. Proliferation and functional maturation of Sertoli cells, and their

- relevance to disorders of testis function in adulthood. *Reproduction*. 2003; 125: 769–784.
76. Berndtson WE, Igboeli G, Parker WG. The Numbers of Sertoli Cells in Mature Holstein Bulls and their Relationship to Quantitative Aspects of Spermatogenesis1. *Biol Reprod*. 1987; 37: 60–67.
 77. Orth JM, Gunsalus GL, Lamperti AA. Evidence From Sertoli Cell-Depleted Rats Indicates That Spermatid Number in Adults Depends on Numbers of Sertoli Cells Produced During Perinatal Development*. *Endocrinology*. 1988; 122: 787–794.
 78. Denis GV, McComb ME, Faller DV, Sinha A, Romesser PB, et al. Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines. *J. Proteome Res*. 2006; 5: 502–511.
 79. Nakamura Y, Umehara T, Nakano K, Jang MK, Shirouzu M, et al. Crystal structure of the human BRD2 bromodomain: insights into dimerization and recognition of acetylated histone H4. *J Biol Chem*. 2007; 282: 4193–4201.
 80. Lamonica JM, Deng W, Kadauke S, Campbell AE, Gamsjaeger R, et al. Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proc. Natl. Acad. Sci. U. S. A*. 2011; 108: E159–E168.
 81. Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, et al. The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. *Mol. Cell*. 2005; 19: 523–534.
 82. Wilson BG, Roberts CWM. SWI/SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer*. 2011; 11: 481–492.
 83. Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet*. 2013; 45: 592–601.
 84. Podcheko A, Northcott P, Bikopoulos G, Lee A, Bommarreddi SR, et al. Identification of a WD40 repeat-containing isoform of PHIP as a novel regulator of beta-cell growth and survival. *Mol. Cell. Biol*. 2007; 27: 6484–6496.
 85. Li S, Francisco AB, Han C, Pattabiraman S, Foote MR, et al.

- The full-length isoform of the mouse pleckstrin homology domain-interacting protein (PHIP) is required for postnatal growth. *FEBS Lett.* 2010; 584: 4121–4127.
86. Kimura J, Nguyen ST, Liu H, Taira N, Miki Y, et al. A functional genome-wide RNAi screen identifies TAF1 as a regulator for apoptosis in response to genotoxic stress. *Nucleic Acids Res.* 2008; 36: 5250–5259.
87. Wang D, Qi H, Zhang H, Zhou W, Li Y, et al. TAF1L promotes development of oral squamous cell carcinoma via decreasing autophagy-dependent apoptosis. *Int. J. Biol. Sci.* 2020; 16: 1180–1193.

Supplementary Materials:

Supplementary Materials can be accessed online at
https://videleaf.com/wp-content/uploads/2022/04/PAGEN2ED-20-14_Supplementary-Materials.zip