Article

Tracing the Origin of the *RSPO2* Long-Hair Allele and Epistatic Interaction between *FGF5* and *RSPO2* in Sapsaree Dog

Mingue Kang [1,†], Byeongyong Ahn [1,†], Seungyeon Youk [1], Yun-Mi Lee [2], Jong-Joo Kim [2], Ji-Hong Ha [3] and Chankyu Park [1,*]

[1] Department of Stem Cell and Regenerative Biotechnology, Konkuk University, Seoul 05029, Korea; mingue5349@gmail.com (M.K.); ahn.b@outlook.com (B.A.); dbrtmddus@hanmail.net (S.Y.)
[2] Department of Biotechnology, Yeungnam University, Gyeongsan 36461, Korea; ymlee@yu.ac.kr (Y.-M.L.); kimjj@ynu.ac.kr (J.-J.K.)
[3] Korean Sapsaree Foundation, Gyeongsan 38412, Korea; jhha@knu.ac.kr
[*] Correspondence: chankyu@konkuk.ac.kr; Tel.: +82-10-8826-1363
[†] These authors contributed equally to this work.

Abstract: Genetic analysis of the hair-length of Sapsaree dogs, a Korean native dog breed, showed a dominant mode of inheritance for long hair. Genome-Wide Association Study (GWAS) analysis and subsequent Mendelian segregation analysis revealed an association between *OXR1*, *RSPO2*, and *PKHD1L1* on chromosome 13 (CFA13). We identified the previously reported 167 bp insertion in *RSPO2* 3' untranslated region as a causative mutation for hair length variations. The analysis of 118 dog breeds and wolves revealed the selection signature on CFA13 in long-haired breeds. Haplotype analysis showed the association of only a few specific haplotypes to the breeds carrying the 167 bp insertion. The genetic diversity in the neighboring region linked to the insertion was higher in Sapsarees than in other Asian and European dog breeds carrying the same variation, suggesting an older history of its insertion in the Sapsaree genome than in that of the other breeds analyzed in this study. Our results show that the *RSPO2* 3' UTR insertion is responsible for not only the furnishing phenotype but also determining the hair length of the entire body depending on the genetic background, suggesting an epistatic interaction between *FGF5* and *RSPO2* influencing the hair-length phenotype in dogs.

Keywords: Sapsaree dogs; hair length; *RSPO2*; *FGF5*; Korean native dogs

1. Introduction

Mammalian hair length is regulated by the activation and inhibition of the Wnt signaling pathway in the hair dermal papilla (DP) and hair matrix, which determine the hair growth cycle divided into anagen, catagen, and telogen [1–5]. In previous studies, the R spondin family, including *R-spondin-2* (*RSPO2*), has been identified to maintain the anagen phase by forming a positive feedback loop of Wnt signaling by inhibiting the Znrf3/Rnf43 complex that antagonizes the Wnt signaling pathway in the hair matrix [5–7]. Additionally, *fibroblast-growth-factor-5* (*FGF5*) induces the catagen phase by terminating the positive feedback loop by suppressing R-spondin expression in DP through the Fgf signaling pathway [5,8,9].

The hair phenotype of modern domestic dog (*Canis lupus familiaris*) breeds is tremendously diverse and is an essential characteristic of each breed. Previous studies have identified mutations in three genes, namely *FGF5*, *RSPO2*, and *keratin-71* (*KRT71*), to be associated with hair length, furnishing, and curl shape, respectively, in dogs [10,11]. The mutations in *FGF5* and *KRT71* were nucleotide substitutions, but the *RSPO2* variant was caused by the 167 bp insertion of a 9-nucleotide (ATAATGAAC) repeat located in the 3' UTR of *RSPO2* transcripts.

Most domestic dog breeds have been developed in the last two hundred years as a result of direct selection designed to meet working or aesthetic requirements [12–15]. The Sapsaree is a medium-sized, long-haired breed of dog native to Korea and frequently appears in the paintings of the Chosun dynasty, the last imperial dynasty of Korea, indicating their presence in the Korean peninsula for at least a few hundred years [16]. Sapsarees were close to extinction in the mid-1980 but were restored from a small number of animals in Korea. During the restoration process, offspring with short-hair phenotypes different from those of the long-haired parents appeared, but the underlying mechanism for hair length variation has not been elucidated. The short-haired Sapsaree has also appeared in the old paintings of the Chosun dynasty (http://www.sapsaree.org/; accessed on 26 September 2021) [16].

In this study, we showed that the hair phenotypes of Sapsarees were caused by the same mutation previously identified in the *RSPO2* gene. We also analyzed selection signatures for the hair length phenotype and analyzed the phylogenetic relationships of the identified haplotypes for the region in domestic dogs. We also investigated the mode of inheritance of *FGF5* and *RSPO2* for hair length phenotypes in domestic dogs.

2. Materials and Methods

2.1. Animals and DNA Preparation

Blood samples and pedigree information of 61 long-haired and 53 short-haired Sapsaree dogs were provided by the Sapsaree Research Foundation in Gyeongsan, Korea. Genomic DNA was isolated from 1 mL of peripheral blood using the QIAamp DNA Blood Midi Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Animal welfare and experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of Konkuk University and conducted according to the accepted protocol.

2.2. Genotyping

Single nucleotide polymorphism (SNP) typing was performed using an Illumina 170 K CanineHD BeadChip (Illumina, San Diego, CA, USA) according to the manufacturer's protocols. Genotype calls were generated using an Illumina GenomeStudio Genotyping Module (Illumina). Pairs of individuals with genetic similarities above 99% were regarded as twins by the identity-by-state test implemented using PLINK v1.07 [17]. Low-quality SNPs with minor allele frequency (MAF) < 0.05% and genotype call rate < 90%, and those that significantly deviated from the Hardy–Weinberg equilibrium ($p < 0.000001$) were removed.

2.3. GWAS and Linkage Analysis

The hair length phenotype (long versus short) data were modeled using a generalized linear mixed model. The logistic regression

$$log\left(\frac{\pi(x)}{1 - \pi(x)}\right) = Xb + SNPp + Zu + e \tag{1}$$

was fitted: b vector containing the fixed effects associated with the traits, while gender was fitted as a fixed effect and age as a covariate; X is the design matrix containing the fixed effects related to the trait, p is the vector of the effect of SNP effect, u represents the vector of the random effects related to the animal effects for the trait, Z is the design matrix that corresponds to the animal effects of the trait, and e is a vector containing the residual. The analysis was conducted using the SAS software version 9.1 (www.sas.com; accessed on 26 September 2021). SNPs linked genomic regions were estimated using PLINK 1.9 (www.cog-genomics.org/plink/1.9/; accessed on 26 September 2021; [18]). The filtering criteria were the same as described above. Linkage disequilibrium (LD) window size was set as 1000 kb, and variants with LD (r^2) > 0.7 were used for linkage analysis.

2.4. Whole-Genome Sequencing Data

The variant call format (VCF) data of 722 dogs was retrieved from NCBI (accession PR-JNA448733, https://www.ncbi.nlm.nih.gov/bioproject/; accessed on 26 September 2021). In addition, the genome sequencing data of 151 dogs, including five breeds of Northeast Asian breeds (PRJNA782070), were retrieved from the NCBI Short Read Archive database, and variant calling was conducted (Table S1). Briefly, whole-genome sequencing data were mapped to the dog reference genome CanFam3.1 (RefSeq accession: GCF_000002285.3) using BWA MEM v0.7.17-r1188 [19]. Base qualities were recalibrated using BaseRecalibrator in the Genome Analysis Toolkit (GATK) package v4.1.9.0 [20]. The joint variant was created using GATK GenotypeGVCFs. Regions of high strand bias (Fisher Strand > 30.0), low quality (QualityOfDepth (QD) < 2), and high complexity with more than three SNPs within a 35 bp window were removed using GATK VariantFiltration. To create a merged VCF file, common SNPs from 151 VCF data, including 5 Sapsarees and 722 VCF data from Plassais et al. [15], were integrated using the -merge option of bcftools 1.13. SNPs were subsequently filtered according to the targeted genomic regions using the thresholds of MAF < 0.01, minimum filtered depth (DP) < 10, and minimum genotype quality (GQ) < 30. Subsequently, individuals with variant call rates under 90% were filtered out. Data from the mixed or unknown breeds were also removed. The same criteria were used for all analyses based on the merged VCF file.

2.5. PCR and Sanger Sequencing

Genotyping PCR was conducted using 15 ng of DNA to determine the presence and absence of the 9-nucleotide repeat insertion in the 3′ UTR of *RSPO2* using previously described primers (forward: 5′-TGGCTAAAGAAAACTTCCACAA-3′; reverse: 5′-AAATTACCATCATGAGACCATGC-3′) [11]. PCR products were analyzed on a 1% agarose gel, and the genotype was determined according to the product sizes (817 and 676 bp for insertion and wildtype alleles, respectively). PCR products were prepared for direct sequencing as described previously [21]. Sequencing reactions were carried out with 30 ng of amplicons from homozygous individuals and 2 pmol of primers identical to PCR amplification for both directions using an ABI PRISM BigDye^{TM} Terminator Cycle Sequencing Kit (Applied Biosystems, Massachusetts, USA) following the manufacturer's protocol. The reactions were analyzed using an ABI 3730 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA). The sequencing chromatograms were analyzed using CLC main workbench 7.8.1 software (CLC bio, Aarhus, Denmark).

2.6. Tajima's D and Nucleotide Diversity

Tajima's D was calculated using the Tajima's D option in vcftools 0.1.17 [22], with a window size of 10,000. Pairwise nucleotide diversity (π) of a genomic region was calculated for breeds with typing data of more than three individuals using vcftools 0.1.17 [22] with "–window-pi 20000," and "–window-pi-step 10000" options.

2.7. Haplotype Analysis

Haplotype blocks were determined using PLINK v1.9 by assigning the case (*RSPO2* 3′UTR insertion homozygotes) and control (*RSPO2* 3′UTR wildtype homozygotes) for each individual. The genotype information of *RSPO2* has been reported previously [11]. The output files were used to define haplotype blocks using Haploview 4.2 [23], and each block was defined by confidence intervals from Gabriel et al. [24] with default options. Haplotype phasing for the estimated haplotype blocks was conducted using BEAGLE v 5.2 [25] with default options. The breeds with SNP typing results from more than three animals were used. For the phased haplotype sequence, a haplotype tree was constructed using Haplotype-Viewer [26], and pairwise distances among haplotypes were estimated using Dnaml [27]. The geographical origin of dog breeds was referred from the Federation Cynologique Internationale (FCI. http://www.fci.be/en/; accessed on 26 September 2021).

2.8. Phylogenetic Analysis and Fixation Index ($F_{ST}{}^{R}$)

Unweighted Pair Group Method with Arithmetic Mean (UPGMA) clustering and neighbor-joining trees were constructed using VCF2poptree [28] with a pairwise diversity output format and visualized using MEGA-X 10.2.2 [29]. Evolutionary distances between individuals were estimated based on the difference in the number of SNPs. $F_{ST}{}^{R}$ was calculated using a published R script [30] using a variant call rate of 0.95, according to the following equation [31]:

$$F_{ST}{}^{R} = \frac{\sum_k \hat{N}^{[k]}}{\sum_k \hat{D}^{[k]}} \tag{2}$$

For $F_{ST}{}^{R}$ estimation, dog breeds with data of more than four individuals were selected.

2.9. Multidimensional Scaling

Multidimensional scaling (MDS) was performed using the informative SNPs in the CFA13:8,530,000-8,730,000 bp region from 363 dogs. A pairwise genetic distance matrix was calculated according to the ratio of the nucleotide difference (p-distance) between individuals using VCF2Dis (https://github.com/BGI-shenzhen/VCF2Dis; accessed on 26 September 2021), and the result was displayed in the MDS plot using the cmdscale function of R (ver 4.1.1; R Core Team, 2021).

3. Results

3.1. Dominant Inheritance of the Long-Hair Phenotype in Sapsarees

The Sapsaree is a Korean native dog breed with medium body size; long hair is its most prominent feature [32]. However, individuals with short hair also appear in the population (Figure S1). To understand the mode of inheritance and the variation in the fur phenotype of Sapsarees, we evaluated the Mendelian segregation of the fur phenotypes of 114 Sapsarees, 61 long-haired and 53 short-haired individuals (Table S2). The pedigree information of the dogs is described in Figure S2. In the population, we identified two crosses with both long-haired parents that produced both long- and short-haired offspring in the same litter. The total number of long- and short-haired offspring from the crosses was two and three, respectively. In contrast, crosses between short-haired parents produced offspring with a phenotype identical to that of the parents in three crosses. The results indicate the dominant Mendelian inheritance of long-hair over the short-hair phenotype, and there was no exception.
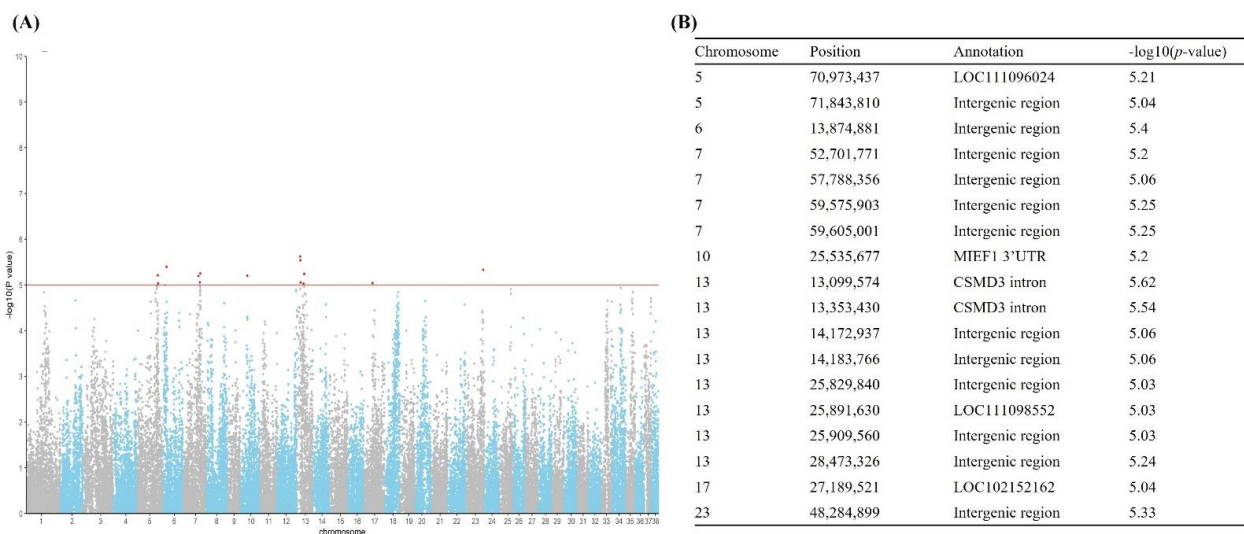
3.2. Mapping the Sapsaree Hair Length Variation-Associated Region to Dog Chromosome 13

Genome-wide SNP typing of 114 Sapsarees using the 170 K canine chip was carried out for a GWAS analysis to identify hair length phenotype-associated regions in the Sapsaree genome. Low-quality SNPs were filtered out, and logistic regression analysis was performed using GLMM. We found that 18 out of 79,229 informative SNPs showed significant associations (p-value $< 10^{-5}$) with the hair length variation of Sapsarees (Table S3). The 18 SNPs were mapped to seven different *Canis familiaris* chromosomes (CFA) with the greatest number of SNPs ($n = 8$) on CFA 13. However, none of the significant SNPs were found in the coding region.

3.3. Identification of RSPO2 as a Candidate for the Hair Length Variation of Sapsarees

The annotation of the eight SNPs on CFA13 showed that one SNP was located in the non-coding RNA (ncRNA) (LOC111098552), two SNPs in the intronic region of *CSMD3*, and the rest in the intergenic region (Figure 1). However, none of the genes near the SNPs were identified as putative candidate genes for the hair phenotype, considering their ontology. Therefore, we manually conducted Mendelian segregation analysis for the genotyping data of CFA 13 using a 170 K canine chip and identified an additional nine SNPs unidentified in the GWAS analysis for their associations with the hair phenotypes (Table S4). We then

conducted a linkage analysis for all the identified SNPs (*n* = 17) to further define the regions of candidate genes (r^2 > 0.7) using PLINK against the VCF of 151 dogs of 20 selected breeds according to geographical origins and hair phenotypes (Table S1: source 1). We found that the regions containing *OXR1*, *RSPO2*, and *PKHD1L1* were strongly associated with the SNPs (Table S5). Interestingly, this result corresponds to previous studies in which the upregulation of *RSPO2* expression by the insertion of a 167 bp segment of a nine-nucleotide repeat (namely, *RSPO2*-repeat) in the 3′ untranslated region (UTR) caused the furnishing phenotype of dogs [11]. Therefore, we assessed the presence of the *RSPO2*-repeat in long-haired Sapsarees by sequence analysis of the *RSPO2* 3′ UTR-specific amplicons (Figure S3) and found *RSPO2*-repeat (nine repeats of ATAATGAAC) in only long-haired Sapsarees.

(A)

(B)

Chromosome	Position	Annotation	-log10(*p*-value)
5	70,973,437	LOC111096024	5.21
5	71,843,810	Intergenic region	5.04
6	13,874,881	Intergenic region	5.4
7	52,701,771	Intergenic region	5.2
7	57,788,356	Intergenic region	5.06
7	59,575,903	Intergenic region	5.25
7	59,605,001	Intergenic region	5.25
10	25,535,677	MIEF1 3'UTR	5.2
13	13,099,574	CSMD3 intron	5.62
13	13,353,430	CSMD3 intron	5.54
13	14,172,937	Intergenic region	5.06
13	14,183,766	Intergenic region	5.06
13	25,829,840	Intergenic region	5.03
13	25,891,630	LOC111098552	5.03
13	25,909,560	Intergenic region	5.03
13	28,473,326	Intergenic region	5.24
17	27,189,521	LOC102152162	5.04
23	48,284,899	Intergenic region	5.33

Figure 1. Hair length phenotype analysis using 114 Sapsarees genotyped on 170 K SNP makers. (**A**) Manhattan plot; (**B**) Maker list showing significance (*p*-value < 10^{-5}).

3.4. Confirmation of the RSPO2-Repeat Insertion as the Underlying Mechanism of Hair Phenotype Variations in Sapsarees using Pedigree Typing

Because the *RSPO2*-repeat was reported as the causative mutation for the furnishing phenotype in dogs, we investigated if the *RSPO2*-repeat was responsible for the phenotype of long-haired Sapsarees using *RSPO2*-repeat typing against the Sapsaree pedigree. The genotyping PCR amplification produced 817 or 676 bp amplicons depending on the presence or absence of the *RSPO2*-repeat, respectively. We confirmed the complete correlation between the long-hair phenotype and the presence of 817 bp products (Figure 2). The 676 bp allele was detected in both long and short-haired Sapsarees, which was consistent with the dominant inheritance of the long-hair allele in Sapsarees. These results indicate that the *RSPO2*-repeat was indeed responsible for the long-hair phenotype in Sapsarees.
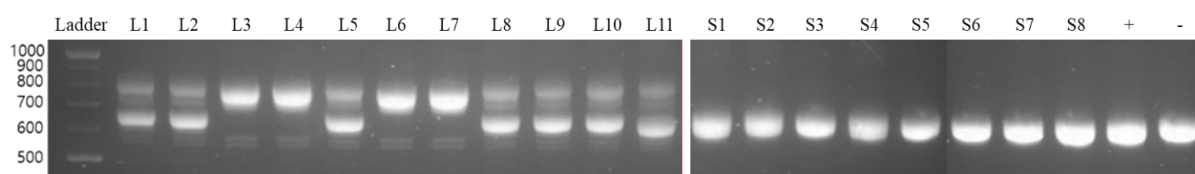

Figure 2. Gel image of the genotyping PCR of *RSPO2* 3′ UTR to differentiate the long- and short-hair allele. Animal ID is shown on top with their phenotypes (L and S for long- and short-hair phenotypes, respectively). The target size of the PCR product for the long- and short-hair alleles was 817 bp and 676 bp, respectively. "+" and "-" indicates positive and negative controls, respectively.

3.5. Selective Sweep of the RSPO2-Associated Regions in Long-Haired Dog Breeds

To evaluate the presence of selection signatures of the *RSPO2* region on CFA 13 (CFA13:8,110,227–9,111,951 bp) among long-haired dog breeds, we analyzed the allele frequency and distribution pattern of 1,213 SNPs identified from a 1 Mb region on CFA 13—500 kb upstream and downstream each of *RSPO2* 3′ UTR—using two different datasets, that is, 18 breeds of 56 *RSPO2*-repeat homozygous individuals and 118 breeds of 369 dogs regardless of *RSPO2* genotype (Table S1). The results showed the complete fixation of an allele in CFA13:8,300,000–8,310,000, 8,500,000–8,510,000, 8,620,000–8,630,000, 8,660,000–8,680,000, and 8,950,000–8,960,000, 8,970,000-8,980,000 bp regions in *RSPO2*-repeat homozygotes (Tables S6 and S7). Selective sweeps with Tajima's D > 95% confidence interval (CI) were identified from the regions, CFA13:8,300,000-8,310,000, 8,500,000–8,700,000, and 8,950,000–9,060,000 bp (Figure 3). The *RSPO2* 3′ UTR was also included in the region of selection sweep. In contrast, Tajima's D values from the dataset, including both long- and short-haired dog breeds, did not display a selection signature for the region.
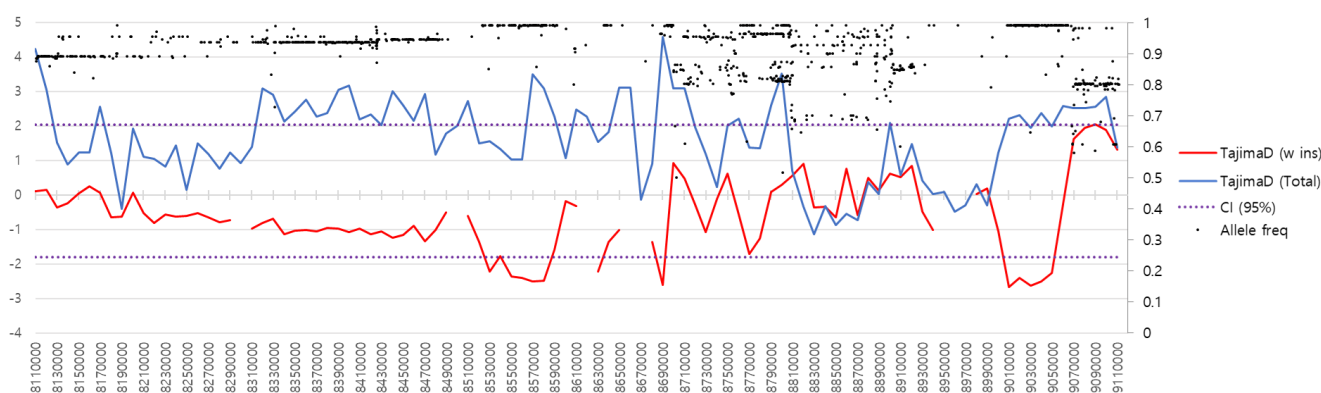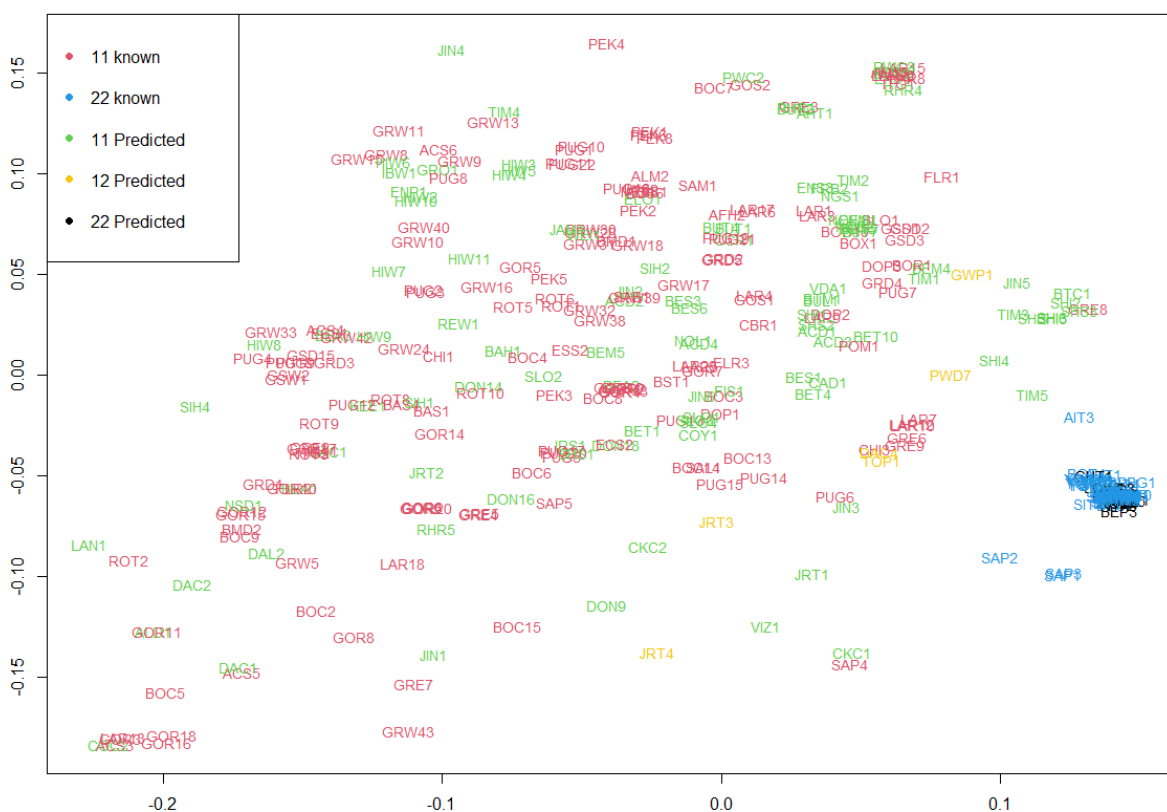

Figure 3. Allele frequency and selection signature for the *RSPO2* 3′ UTR-linked genomic region on CFA 13. The allele frequencies of 1,225 SNPs in the CFA13:8,110,000−9,110,000 bp region of 56 *RSPO2* 3′ UTR insertion homozygous dogs are indicated by black dots. Tajima's D is depicted as red and blue lines for 56 dogs homozygous for the *RSPO2* 3′ UTR insertion and 369 dogs consisting of all genotypes. Dotted lines indicate the 95% confidence interval (CI) of Tajima's D value according to the beta distribution corresponding to *n* = 55.

3.6. Common Origin of the RSPO2 Allele among Diverse Long-Hair Dog Breeds

To understand the ancestry of the *RSPO2*-repeat allele in long-haired dog breeds, phylogenetic analysis was carried out using 770 SNPs identified from a 200 kb region surrounding the *RSPO2* 3′ UTR (CFA13:8,530,000−8,730,000 bp) from 92 individuals of 24 breeds selected based on their geological origins (Figure 4, Table S1, and Figure S4). In both neighbor-joining and UPGMA trees, the individuals of long-haired breeds, including the Shih Tzu (SIT), Tibetan Terrier (TIT), Portuguese Water Dog (PWD), Yorkshire Terrier (YOT), and long-haired Sapsarees (SAP), were clustered close (Figure 4 and Figure S4). Chinese crested dogs (CHC) were also clustered within the long-haired dog cluster, suggesting the presence of the *RSPO2*-repeat in the breed and additional changes in hair phenotypes due to the influence of other mutations on their hairless body with a crested hair phenotype. The average genetic distance among the long-haired Sapsarees (SAP1, 2, and 3) based on the number of nucleotide differences (*n* = 22.67) was much larger than the average genetic distance between the other breeds in the cluster (*n* = 13.93), suggesting a possibly longer history of the long-hair phenotype of Sapsarees (Table S8). Consistently, Sapsarees were shown as an outgroup within the long-haired dog cluster in the tree (Figure 4 and Figure S4). Homozygous short-haired Sapsarees (SAP4 and 5) formed clusters with short-haired breeds native to China and Korea, such as the Pug (PUG), Jindo (JIN), and DongGyeongi (DON) and other European breeds like the Golden Retriever (GOR), and Australian breed Australian Cattle dog (ACD). Gray wolf (GRW) and Himalayan wolf

(*Canis lupus chanco*) (HIW) constituted a separate cluster. Phylogenetic relationships of the *RSPO2* region from short-haired dog breeds were more diverse and distantly related than those of the long-haired breeds. The MDS plot using the pairwise distance matrix generated from the CFA 13: 8,530,000−8,730,000 bp region, identical to the phylogenetic analysis from 363 dogs (Table S1), also showed the clustering of the long-haired breeds with the *RSPO2*-repeat allele (Figure 5). These results suggest that the allelic origin of the *RSPO2* gene in most long-haired dog breeds is common.

Figure 4. A phylogenetic tree for the 200 kb region on dog chromosome 13 with selection sweep for the hair length phenotype. A UPGMA phylogenetic tree was constructed using 785 SNPs from a total of 92 dogs belonging to 24 breeds with various hair length phenotypes. The branch length was determined on the basis of pairwise nucleotide differences in the CFA13:8,530,000−8,730,000 bp region among individuals. Homozygotes for the *RSPO2* 3' UTR insertion are indicated by closed circles, and animals without the insertion are indicated by open circles. Gray circles indicate heterozygotes. Five Sapsarees were additionally marked with black squares. GOR (Golden Retriever), ACD (Australian Cattle Dog), AFH (Afghan Hound), YOT (Yorkshire Terrier), BOC (Border Collie), BUT (Bull Terrier), GRD (Great Dane), CHC (Chinese Crested), COC (Chow Chow), PWD (Portuguese Water Dog), RHR (Rhodesian Ridgeback), SIH (Siberian Husky), SAM (Samoyed), DON (DongGyeongi), GRW (Grey Wolf), HIW (Himalayan Wolf), JIN (Jindo), PEK (Pekingese), PUG (Pug), SAP (Sapsaree), SHI (Shiba Inu), SIT (Shih Tzu), TIM (Tibetan Mastiff), TIT (Tibetan Terrier).

Figure 5. Multidimensional scaling plot of dog breeds with various hair lengths according to SNP variations. The plot was created on the basis of pairwise distances calculated from SNPs of the haplotype block 20 (CFA13:8,530,000−8,730,000 bp) linked to the *RSPO2* 3' UTR region. The genotypes of *RSPO2* 3' UTR in the top left of the plot were described as follows: "11" for homozygotes of the *RSPO2*-wildtype allele, "22" for homozygotes of the *RSPO2*-repeat allele, and "12" for heterozygotes. The words "known" and "predicted" indicate the use of available data (Cadieu et al. 2009) and prediction using the linkage analysis in this study, respectively. For information on the abbreviation of each breed, refer to Table S1.

3.7. Haplotype Diversity of the Region Linked to RSPO2-Repeat in Domestic Dogs

The haplotype block associated with the *RSPO2*-repeat was estimated for the 200 kb region (CFA13:8,530,000−8,730,000 bp), which showed a selection signature. A total of 224 individuals−55 belonging to 19 breeds homozygous for *RSPO2*-repeat and 169 belonging to 35 breeds homozygous for *RSPO2*-wildtype allele−were used to define the haplotype blocks associated with the *RSPO2*-repeat allele (Table S1). Among the 29 haplotype blocks defined, haplotype block 20 was the largest and consisted of 57 SNPs (Table S9). For haplotype block 20, we identified 12 major haplotypes with a haplotype frequency >1% (Figure S5) that could be differentiated using 11 tag SNPs. Interestingly, a haplotype consisting of the SNPsT/G/T/A/G/T/T/A/A/G/A/A/G/C/T/T/G/C/A/A/ A/G/T/G/A/G/G/G/A/A/T/G/T/C/T/A/C/G/T/A/T/G/G/A/A/A/C/T/C/G/ C/T/G/C/A/T/C was highly specific to the breeds carrying the *RSPO2*-repeat allele, with a few exceptions (*p*-value = 7.13×10^{-89}, Table S9). Subsequently, an additional haplotype analysis using the entire dataset of 363 individuals including heterozygotes (*n* = 13), genotype-undefined individuals (*n* = 126), and homozygotes (*n* = 224), was conducted using the 57 SNPs, and the 12 major haplotypes (HapA to L) were defined (Figure 6). HapA was distributed across the largest number of breeds (*n* = 60). Hap J was the most distant from the other haplotypes. HapD was associated with the *RSPO2*-repeat insertion. The distribution of haplotypes in block 20 did not have any geological associations.
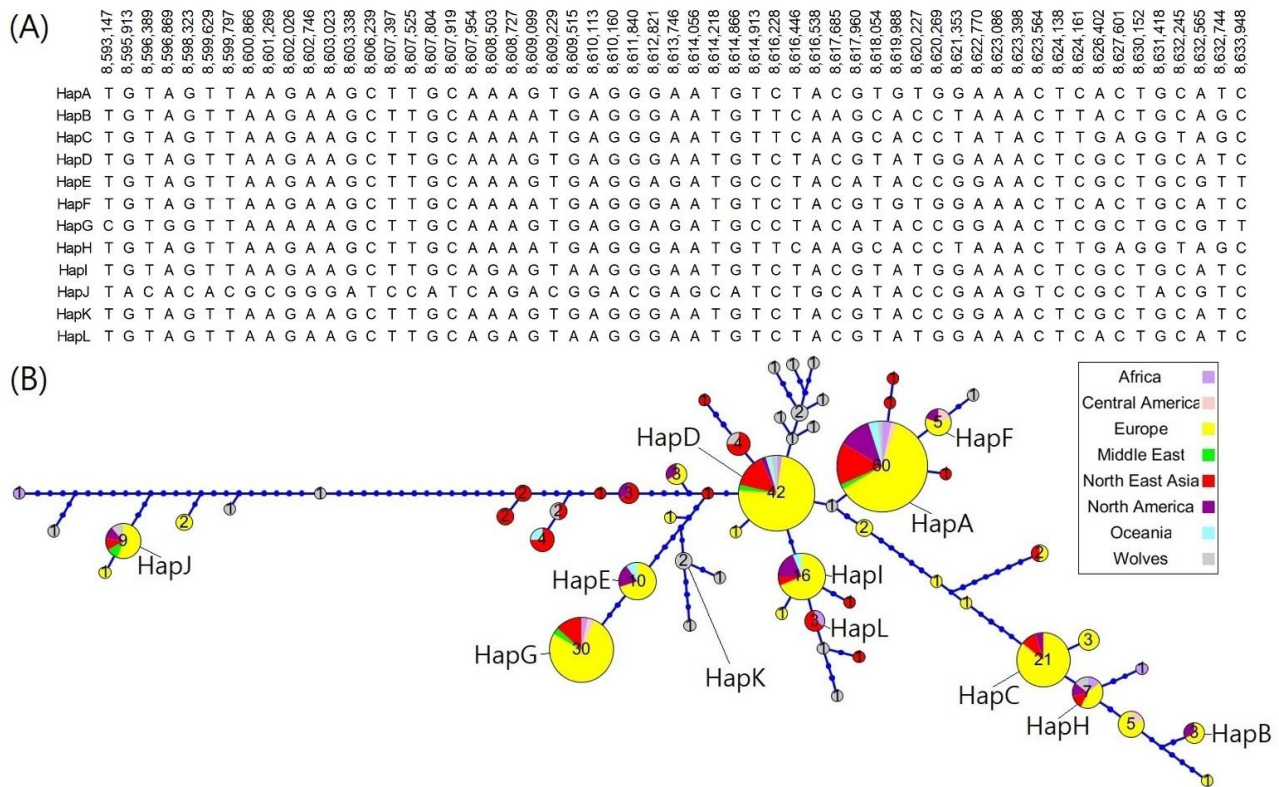
Figure 6. Haplotype tree showing the relationships of the 58 haplotypes belonging to the haplotype block 20 of the *RSPO2* 3′ UTR-linked region on dog chromosome 13. The 12 major haplotypes (HapA to L) with frequencies >0.01 are shown with the corresponding SNPs and chromosomal coordinates in (**A**) and indicated within or next to each haplotype node in the haplotype tree (**B**). Haplotype nodes are colored differently according to the geological origin of the breed, and the number of breeds with the corresponding haplotype is specified at each node. The distance between the nodes was determined based on nucleotide differences among haplotype sequences, and each node represents a difference of 1 SNP from the adjacent nodes or dots. Blue dots indicate intermediate sequence states between two different haplotype node sequences.

Because the complete differentiation of the *RSPO2*-repeat allele from the wildtype allele was not achieved by haplotype analysis using 57 SNPs, we conducted an additional haplotype phasing using all SNPs (*n* = 178) present in the haplotype block 20, which increased the number of haplotypes to 93 (Figure S6). Finally, five haplotypes were identified to be specific to the *RSPO2*-repeat allele-carrying dog breeds. In addition, five haplotypes (Hap1 to 5) shared two common SNPs at positions 8,610,423 (T) and 8,610,424 (A), in contrast to G and C, respectively, of the remaining haplotypes corresponding to short-haired breeds. Consequently, the occurrence of 93 haplotypes divided the animals into 278 *RSPO2*-wildtype homozygotes, 77 *RSPO2*-repeat homozygotes, and 8 *RSPO2* heterozygotes (Tables S1 and S10) and showed complete agreement with the reported genotypes for each breed [11]. Therefore, we suggest that the two SNPs can be used to predict the allelic type of *RSPO2* without the *RSPO2* 3′ UTR region. The five haplotypes associated with the *RSPO2*-repeat were closely clustered in the phylogenetic tree (Figure S7), which was in line with the clustering pattern in the MDS plot (Figure 5).

3.8. Genetic Diversity in the RSPO2-Linked Region in Sapsarees

We compared nucleotide diversity (π) of the genomic region linked to *RSPO2* (CFA13:8, 530,000–8,730,000 bp) in 57 *RSPO2*-repeat homozygotes belonging to 13 breeds with at least three animals per breed (Table S1). The mean π in a 20 kb window was estimated to be 0.000034 (SD = 0.000054). Interestingly, long-haired Sapsarees showed the most

nucleotide diversity ($\pi = 0.00018$) (Table S11), which is consistent with the results of the phylogenetic analysis in Figure 4. We also analyzed breed differentiation by estimating F_{ST}^R for the CFA13:8,530,000–8,730,000 bp region between the breeds, each with at least four individuals (Table S1). The pairwise F_{ST}^R for the 200 kb of the *RSPO2*-linked region between Sapsarees and other breeds was lowest with Jack Russel Terrier (JRT) and highest with German Shepherd Dog (GSD). Northeast Asian dog breeds with relatively low F_{ST}^R with Sapsarees were Jindo, DongGyeongi, and Shiba Inu (Figure 7), consistent with their geographical origins.
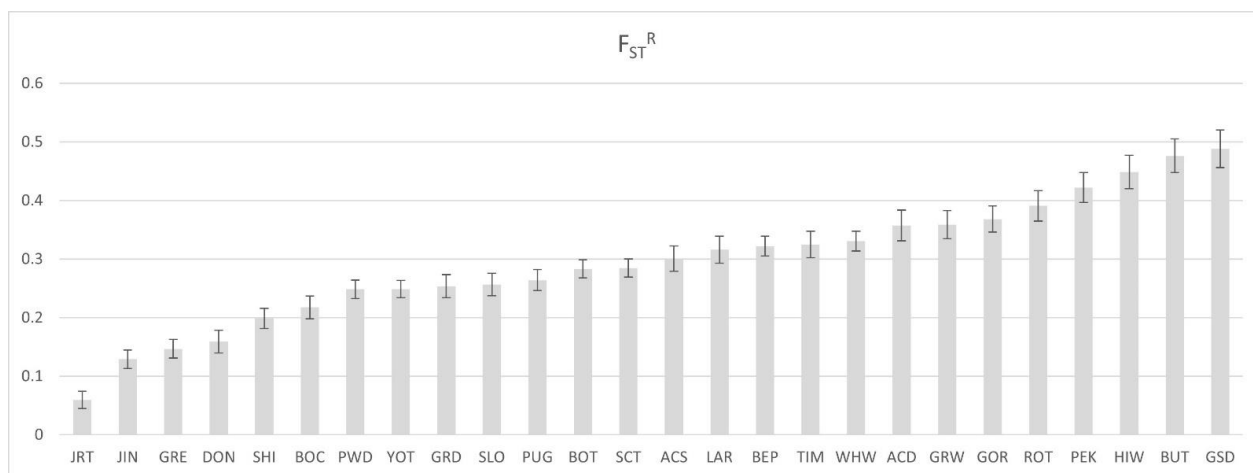
Figure 7. Pairwise F_{ST}^R calculated from the *RSPO2* 3' UTR-associated region of the long- and short-haired Sapsarees and other dog breeds. The pairwise F_{ST}^R between the breeds was calculated based on the definition of Reich et al. (2009). The error bar at the top of each bar graph indicates the 95% confidence interval.

4. Discussion

Native breeds of domestic animals are valuable in helping unravel the molecular mechanisms underlying their diverse phenotypic characteristics obtained through their adaptation to natural or human-induced change. Several GWAS studies have identified genes or loci significantly associated with diverse morphological traits in dogs and highlighted the molecular mechanisms and importance of genetic variations in phenotypic diversity [15,33,34]. In this study, we determined the causative gene for hair length variations in Sapsarees, a native dog breed in the Korean peninsula, and investigated the allelic origin, selection signature, and haplotypic variations of the neighboring sequence of *RSPO2* across diverse breeds ($n = 172$) of domestic dogs (Table S1).

Our initial GWAS analysis mapped the region of association to the long-hair phenotype of Sapsarees on CFA 13, but with less precision. This may be due to the dominance of the long-hair allele and the small sample size ($n = 114$). Probing several other chromosomal regions showing significance also did not lead to the identification of any hair growth-related genes. However, the presence of multiple markers on CFA13 showing strong associations with the long-hair phenotype was critical for the identification of the causal variants underlying hair length variation in Sapsarees.

Sapsarees, a native dog breed, have existed in the Korean peninsula for at least a few hundred years and carry an identical allele responsible for the hair phenotype of the European dog breeds [35] despite distant genetic relationships between Sapsarees and Western dog breeds [36–38]. Currently, 197 long-haired dog breeds have been officially recognized by the American Kennel Club (AKC, www.ack.org/dog-breeds; accessed on 26 September 2021), and their hair length variations have been explained by the *FGF5* and *RSPO2* variations [11,39]. The presence of the identical *RSPO2*-repeat allele among the *RSPO2*-associated long-haired dog breeds, including Sapsarees, suggests a common origin of the allele or an ancestral effect in long-haired dog breeds. This suggests that the genetic

origin of hair length phenotypes could be different from the genetic background estimated from whole genomes [13,40–44].

The presence of both *RSPO2*-repeat and wildtype alleles in a single breed has been reported on several occasions, including in the Portuguese Water Dog, Brussels Griffon, Dachshund, and Sapsarees [11,35]. The maintenance of other phenotypic characteristics, except for hair length in a breed, could be explained by the occurrence of genetic mutations in hair length association genes. However, it is less likely to carry the same causative allele across several breeds through independent mutational events. For example, five recessive variants of *FGF5* have been reported as causal mutations for the long-hair phenotype in diverse dog breeds with the *FGF5*:p.Cys95Phe mutation as the most common among the breeds [39]. In contrast, an identical *RSPO2* variant has been found in all the breeds analyzed to date, however, the repeat number of the nine-nucleotide repeat in the *RSPO2* 3′ UTR insertion allele differs slightly in some breeds [35]. Perhaps missense mutations in *RSPO2* could be harmful, considering its role in Wnt signaling, which is critical for animal development [45–47].

To understand the possible mechanisms involved in the insertion of the nine-nucleotide repeat in *RSPO2* 3′ UTR, we blasted 150 bp each of the upstream and downstream sequence of the 3′ insertion junction of *RSPO2* as a query against the reference genome of dogs using NCBI nucleotide BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi; accessed on 26 September 2021). Although the results showed several matches on chromosomes other than CFA 13, we were unable to draw any conclusion on the mechanism of simple repeat insertion.

A previous study reported an *RSPO2*-associated region showing a high F_{ST} [48]. However, the region is less directly related to the 200 kb genomic region strongly linked to the *RSPO2*-repeat allele in Sapsarees. The higher genetic diversity in the 200 kb genomic region linked to the *RSPO2*-repeat allele in Sapsarees than the other dog breeds suggests that the age of the *RSPO2* 3′ UTR insertion in Sapsarees may be older or less affected by intensive selection than in the other breeds analyzed in this study despite the bottleneck effect of breed rebuilding. However, the haplotype diversity in the neighboring region of the *RSPO2*-wildtype allele in 119 dog breeds showed much more diversity (Hap6–93) than that in the neighboring region of the *RSPO2*-repeat allele (Hap1–5) (Figure S7). Unexpectedly, the degree of population differentiation (F_{ST}^{R}) against the sequences of the 200 kb region of Sapsarees was the lowest (0.0594) with Jack Russel Terrier, a European breed, followed by other Korean native breeds, namely Jindo and DongGyeongi (Figure 7). This may suggest a common origin of the *RSPO2*-repeat allele in the domestic dog breeds.

In previous studies, the *RSPO2*-repeat allele has been described as a causal mutation for the furnishing phenotype. However, the combined genotype of *FGF5* and *RSPO2*, in which the presence and absence of mutant alleles were described as "+" and "-", respectively, showed that the hair length of Sapsarees was also significantly affected by the *RSPO2*-repeat allele (Table 1). Among Northeast Asian dogs, the long-haired Sapsarees (*FGF5* +, *RSPO2* +) showed allelic combinations of the two genes identical to those of Shih Tzu and Tibetan Terrier. This pattern also appeared in European breeds such as the Old English Sheepdog, Maltese, and Portuguese Water Dog [11]. In contrast, short-haired Sapsarees had the same allelic combination as that of other Northeast Asian breeds, including Pekingese and Tibetan Mastiff, as well as European breeds such as the Newfoundland, Golden Retriever, and English Cocker Spaniel [11]. However, none of the breeds with *FGF5* (−) and *RSPO2* (+) genotype showed the long-hair phenotype. The hair length of the short-haired Sapsarees (*FGF5* +, *RSPO2* −) was significantly shorter than that of other breeds with the same genotype, such as Pekingese and Tibetan Mastiffs. Such phenotypic variation could be due to the interaction of *RSPO2* with other hair cycle regulatory genes or the regulatory sequence variation of *RSPO2*. These results suggest that hair length in dogs is regulated by the epistatic interaction between *FGF5* and *RSPO2*.

Table 1. Hair length phenotypes and the presence of *FGF5* and *RSPO2* variants in Northeast Asian breeds and two selected European breeds.

Breed	Origin	Hair Length	Furnishing	*FGF5*	*RSPO2*
DongGyeongi	Northeast Asia	Short	No	−	−
Jindo	Northeast Asia	Short	No	−	−
Pekingese	Northeast Asia	Long	No	+	−
Pug	Northeast Asia	Short	No	−	−
Sapsaree-long	Northeast Asia	Long	Yes	+	+
Sapsaree-short	Northeast Asia	Short	No	+	-
Shiba Inu	Northeast Asia	Short	No	−	−
Shih Tzu	Northeast Asia	Long	Yes	+	+
Tibetan Mastiff	Northeast Asia	Medium	No	+	−
Tibetan Terrier	Northeast Asia	Long	Yes	+	+
Petit Basset Griffon Vendeen*	Europe	Short	Yes	−	+
Irish Wolfhound*	Europe	Short	Yes	−	+

Note: Plus (+) and minus signs (−) indicate the presence and absence of the long-hair variant allele, respectively. The allele information and hair phenotypes of breeds marked with "*" is from Cadieu et al. (2009).

In this study, we identified *RSPO2* as the causative mutation determining the short hair length phenotype of Sapsarees and provided evidence for the role of the epistatic interaction between *RSPO2* and *FGF5* in hair length phenotype in domestic dogs. We also showed the presence of the *RSPO2* variant-associated selection sweep and haplotypic diversity of the region. The presence of a single *RSPO2* variant in all dog breeds strongly suggests a common origin of the mutation. The significantly higher genetic diversity within the *RSPO2*-repeat allele-linked haplotype block in Sapsarees than in the other long-haired dog breeds may suggest that the integration of the *RSPO2* variant in the genome of Sapsarees happened earlier than in that of other long-hair dog breeds carrying the variant. However, the difference in breeding history, including the intensity of inbreeding, should also be considered. Previous studies have suggested the *RSPO2*-repeat allele to be responsible for the long-hair phenotype in various dog breeds. However, this study further explored the phylogenetics of the *RSPO2* mutation in domestic dogs and its genetic interactions by also analyzing the genetic data of previously undescribed breeds.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/genes13010102/s1, Figure S1: Pictures of native Korean Sapsaree dogs with different hair length phenotypes. In the picture, from left to right, are long- and short-haired Sapsaree dogs. Pictures were from the Korean Sapsaree foundation (http://eng.sapsaree.org/ accessed on 26 September 2021). Figure S2: Pedigree data of 114 Sapsarees used in this study, Figure S3: The nucleotide sequence alignment of the *RSPO2* 3′ UTR region between the long- and short-haired variants of Sapsarees, Figure S4: Neighbor-joining tree constructed using the CFA13:8,530,000–8,730,000 bp region of 92 dogs belonging to 24 dog breeds with various hair length phenotypes, Figure S5: Haplotype block 20 estimated by linkage disequilibrium of SNPs present in the *RSPO2* 3′ UTR linked region on dog chromosome 13, Figure S6: Result of multiple sequence alignment using 93 haplotype sequences of the *RSPO2* 3′ UTR selective sweep region, Figure S7: Phylogenetic relationships among the haplotypes of the haplotype block 20 region, Table S1: Details of the 795 dogs and their genome sequences used in the study, Table S2: List of the hair-length data of the 114 Sapsaree dogs used in this study, Table S3: Results of GWAS analysis from the typing results of 114 Sapsarees genotyped on 170K CanineHD BeadChip, Table S4: List of SNPs on CFA13 showing complete Mendelian segregation with the hair length phenotype of Sapsarees, Table S5: List of SNPs and linked genes associated with hair-length phenotype, Table S6: Estimated Tajima's D for the hair length phenotype-associated region of chromosome 13 (CFA13: 8,110,000–9,120,000 bp), Table S7: Allelic distribution of SNPs in the *RSPO2*-linked region of 56 dogs, Table S8: Estimated genetic distances of dogs belonging to the 6 long-haired breeds selected using the pairwise comparison of the 200 kb *RSPO2* 3′ UTR insertion-linked region, Table S9: Results of haplotype block estimation using the HaploView program, Table S10: Prediction of *RSPO2* 3′ UTR insertion genotypes according to the haplotypes of the *RSPO2* 3′ UTR-linked region for the individuals lacking *RSPO2* sequences, Table S11: Nucleotide diversity

(π) of the *RSPO2* 3′ UTR insertion-linked genomic region of 13 dog breeds homozygous for *RSPO2* 3′UTR insertion.

References

1. Chase, H.B. Growth of the hair. *Physiol. Rev.* **1954**, *34*, 113–126. [CrossRef] [PubMed]
2. Enshell-Seijffers, D.; Lindon, C.; Kashiwagi, M.; Morgan, B.A. Beta-catenin activity in the dermal papilla regulates morphogenesis and regeneration of hair. *Dev. Cell* **2010**, *18*, 633–642. [CrossRef] [PubMed]
3. Hardy, M.H. The secret life of the hair follicle. *Trends Genet.* **1992**, *8*, 55–61. [CrossRef]
4. Kishimoto, J.; Burgeson, R.E.; Morgan, B.A. Wnt signaling maintains the hair-inducing activity of the dermal papilla. *Genes Dev.* **2000**, *14*, 1181–1185. [CrossRef] [PubMed]
5. Harshuk-Shabso, S.; Dressler, H.; Niehrs, C.; Aamar, E.; Enshell-Seijffers, D. Fgf and Wnt signaling interaction in the mesenchymal niche regulates the murine hair cycle clock. *Nat. Commun.* **2020**, *11*, 5114. [CrossRef]
6. Hao, H.X.; Xie, Y.; Zhang, Y.; Charlat, O.; Oster, E.; Avello, M.; Lei, H.; Mickanin, C.; Liu, D.; Ruffner, H.; et al. ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. *Nature* **2012**, *485*, 195–200. [CrossRef]
7. Koo, B.K.; Spit, M.; Jordens, I.; Low, T.Y.; Stange, D.E.; van de Wetering, M.; van Es, J.H.; Mohammed, S.; Heck, A.J.; Maurice, M.M.; et al. Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **2012**, *488*, 665–669. [CrossRef]
8. Hébert, J.M.; Rosenquist, T.; Götz, J.; Martin, G.R. FGF5 as a regulator of the hair growth cycle: Evidence from targeted and spontaneous mutations. *Cell* **1994**, *78*, 1017–1025. [CrossRef]
9. Suzuki, S.; Ota, Y.; Ozawa, K.; Imamura, T. Dual-mode regulation of hair growth cycle by two Fgf-5 gene products. *J. Investig. Dermatol.* **2000**, *114*, 456–463. [CrossRef]
10. Housley, D.J.; Venta, P.J. The long and the short of it: Evidence that FGF5 is a major determinant of canine 'hair'-itability. *Anim. Genet.* **2006**, *37*, 309–315. [CrossRef]
11. Cadieu, E.; Neff, M.W.; Quignon, P.; Walsh, K.; Chase, K.; Parker, H.G.; Vonholdt, B.M.; Rhue, A.; Boyko, A.; Byers, A.; et al. Coat variation in the domestic dog is governed by variants in three genes. *Science* **2009**, *326*, 150–153. [CrossRef]
12. Boyko, A.R.; Quignon, P.; Li, L.; Schoenebeck, J.J.; Degenhardt, J.D.; Lohmueller, K.E.; Zhao, K.; Brisbin, A.; Parker, H.G.; vonHoldt, B.M.; et al. A simple genetic architecture underlies morphological variation in dogs. *PLoS Biol.* **2010**, *8*, e1000451. [CrossRef]
13. Von Holdt, B.M.; Pollinger, J.P.; Lohmueller, K.E.; Han, E.; Parker, H.G.; Quignon, P.; Degenhardt, J.D.; Boyko, A.R.; Earl, D.A.; Auton, A.; et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **2010**, *464*, 898–902. [CrossRef]
14. Ostrander, E.A.; Wayne, R.K.; Freedman, A.H.; Davis, B.W. Demographic history, selection and functional diversity of the canine genome. *Nat. Rev. Genet.* **2017**, *18*, 705–720. [CrossRef]
15. Plassais, J.; Kim, J.; Davis, B.W.; Karyadi, D.M.; Hogan, A.N.; Harris, A.C.; Decker, B.; Parker, H.G.; Ostrander, E.A. Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat. Commun.* **2019**, *10*, 1489. [CrossRef] [PubMed]
16. Ha, J.-H. *Dogs in Korea: An Uncomfortable Truth of Korean Dogs*, 1st ed.; Global Contents: Seoul, Korea, 2017; Volume 1, pp. 1–219.
17. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Sham, P.C. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
18. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* **2015**, *4*, 7. [CrossRef] [PubMed]
19. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]

20. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; Del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **2013**, *43*, 11. [CrossRef]

21. Dinka, H.; Le, M.T.; Ha, H.; Cho, H.; Choi, M.K.; Choi, H.; Kim, J.H.; Soundarajan, N.; Park, J.K.; Park, C. Analysis of the vomeronasal receptor repertoire, expression and allelic diversity in swine. *Genomics* **2016**, *107*, 208–215. [CrossRef]

22. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef] [PubMed]

23. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [CrossRef] [PubMed]

24. Gabriel, S.B.; Schaffner, S.F.; Nguyen, H.; Moore, J.M.; 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*, 2225–2229. [CrossRef]

25. Browning, B.L.; Tian, X.; Zhou, Y.; Browning, S.R. Fast two-stage phasing of large-scale sequence data. *Am. J. Hum. Genet.* **2021**, *108*, 1880–1890. [CrossRef] [PubMed]

26. Salzburger, W.; Ewing, G.B.; Haeseler, A.V. The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Mol. Ecol.* **2011**, *20*, 1952–1963. [CrossRef] [PubMed]

27. Felsenstein, J. *PHYLIP (Phylogeny Interference Package), version 3.7a*; University of Washington: Seattle, WA, USA, 2009.

28. Subramanian, S.; Ramasamy, U.; Chen, D. VCF2PopTree: A client-side software to construct population phylogeny from genome-wide SNPs. *PeerJ* **2019**, *7*, e8213. [CrossRef]

29. Kumar, C.; Song, S.; Jiang, L.; He, X.; Zhao, Q.; Pu, Y.; Malhi, K.K.; Kamboh, A.A.; Ma, Y. Sequence Characterization of DSG3 Gene to Know Its Role in High-Altitude Hypoxia Adaptation in the Chinese Cashmere Goat. *Front. Genet.* **2018**, *9*, 553. [CrossRef]

30. Di Gaetano, C.; Fiorito, G.; Ortu, M.F.; Rosa, F.; Guarrera, S.; Pardini, B.; Cusi, D.; Frau, F.; Barlassina, C.; Troffa, C.; et al. Sardinians genetic background explained by runs of homozygosity and genomic regions under positive selection. *PLoS ONE* **2014**, *9*, e91237. [CrossRef]

31. Reich, D.; Thangaraj, K.; Patterson, N.; Price, A.L.; Singh, L. Reconstructing Indian population history. *Nature* **2009**, *461*, 489–494. [CrossRef]

32. Han, K.-I.; Alam, M.; Lee, Y.-M.; Lee, D.-H.; Ha, J.-H.; Kim, J.-J. A Study on Morphology and Behavior of the Sapsaree: A Korean native dog (Canis familiaris). *J. Anim. Sci. Technol.* **2010**, *52*, 481–490. [CrossRef]

33. Hayward, J.J.; Castelhano, M.G.; Oliveira, K.C.; Corey, E.; Balkman, C.; Baxter, T.L.; Casal, M.L.; Center, S.A.; Fang, M.; Garrison, S.J.; et al. Complex disease and phenotype mapping in the domestic dog. *Nat. Commun* **2016**, *7*, 10460. [CrossRef] [PubMed]

34. Schoenebeck, J.J.; Ostrander, E.A. Insights into morphology and disease from the dog genome project. *Annu. Rev. Cell. Dev. Biol.* **2014**, *30*, 535–560. [CrossRef] [PubMed]

35. Parker, H.G.; Chase, K.; Cadieu, E.; Lark, K.G.; Ostrander, E.A. An Insertion in the RSPO2 Gene Correlates with Improper Coat in the Portuguese Water Dog. *J. Hered.* **2010**, *101*, 612–617. [CrossRef] [PubMed]

36. Gajaweera, C.; Kang, J.M.; Lee, D.H.; Lee, S.H.; Kim, Y.K.; Wijayananda, H.I.; Kim, J.J.; Ha, J.H.; Choi, B.H.; Lee, S.H. Genetic diversity and population structure of the Sapsaree, a native Korean dog breed. *BMC Genet.* **2019**, *20*, 1–11. [CrossRef] [PubMed]

37. Jeong, H.; Choi, B.-H.; Eo, J.; Kwon, Y.-J.; Lee, H.-E.; Choi, Y.; Gim, J.-A.; Kim, T.-H.; Seong, H.-H.; Lee, D.-H.; et al. Statistical analysis and genetic diversity of three dog breeds using simple sequence repeats. *Genes Genom.* **2014**, *36*, 883–889. [CrossRef]

38. Kang, M.; Ahn, B.; Youk, S.; Cho, H.-s.; Choi, M.; Hong, K.; Do, J.T.; Song, H.; Jiang, H.; Kennedy, L.J.; et al. High Allelic Diversity of Dog Leukocyte Antigen Class II in East Asian Dogs: Identification of New Alleles and Haplotypes. *J. Mamm. Evol.* **2021**, *28*, 773–784. [CrossRef]

39. Dierks, C.; Momke, S.; Philipp, U.; Distl, O. Allelic heterogeneity of FGF5 mutations causes the long-hair phenotype in dogs. *Anim. Genet.* **2013**, *44*, 425–431. [CrossRef]

40. Shannon, L.M.; Boyko, R.H.; Castelhano, M.; Corey, E.; Hayward, J.J.; McLean, C.; White, M.E.; Said, M.A.; Anita, B.A.; Bondjengo, N.I.; et al. Genetic structure in village dogs reveals a Central Asian domestication origin. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13639–13644. [CrossRef]

41. Bergstrom, A.; Frantz, L.; Schmidt, R.; Ersmark, E.; Lebrasseur, O.; Girdland-Flink, L.; Lin, A.T.; Stora, J.; Sjogren, K.G.; Anthony, D.; et al. Origins and genetic legacy of prehistoric dogs. *Science* **2020**, *370*, 557–564. [CrossRef]

42. Skoglund, P.; Ersmark, E.; Palkopoulou, E.; Dalen, L. Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Curr. Biol.* **2015**, *25*, 1515–1519. [CrossRef]

43. Wang, G.D.; Zhai, W.; Yang, H.C.; Wang, L.; Zhong, L.; Liu, Y.H.; Fan, R.X.; Yin, T.T.; Zhu, C.L.; Poyarkov, A.D.; et al. Out of southern East Asia: The natural history of domestic dogs across the world. *Cell Res.* **2016**, *26*, 21–33. [CrossRef] [PubMed]

44. Frantz, L.A.; Mullin, V.E.; Pionnier-Capitan, M.; Lebrasseur, O.; Ollivier, M.; Perri, A.; Linderholm, A.; Mattiangeli, V.; Teasdale, M.D.; Dimopoulos, E.A.; et al. Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* **2016**, *352*, 1228–1231. [CrossRef] [PubMed]

45. Knight, M.N.; Hankenson, K.D. R-spondins: Novel matricellular regulators of the skeleton. *Matrix Biol.* **2014**, *37*, 157–161. [CrossRef] [PubMed]

46. Knight, M.N. R-Spondin-2 Modulates Osteoblastogenesis and Bone Accrual Through WNT Signaling. Ph.D. Thesis, University of Pennsylvania, Philadelphia, PA, USA, 2017.

47. Becker, D.; Weikard, R.; Schulze, C.; Wohlsein, P.; Kuhn, C. A 50-kb deletion disrupting the RSPO2 gene is associated with tetradysmelia in Holstein Friesian cattle. *Genet. Sel. Evol.* **2020**, *52*, 1–13. [CrossRef] [PubMed]
48. Vaysse, A.; Ratnakumar, A.; Derrien, T.; Axelsson, E.; Pielberg, G.R.; Sigurdsson, S.; Fall, T.; Seppala, E.H.; Hansen, M.S.; Lawley, C.T.; et al. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet.* **2011**, *7*, e1002316. [CrossRef] [PubMed]