Journal of
Animal Breeding and Genetics

J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

A genome-wide association study for canine cryptorchidism in Siberian Huskies

X. Zhao, S. Onteru, M. Saatchi, D. Garrick & M. Rothschild

Department of Animal Science, Iowa State University, Ames, IA, USA

Keywords
Siberian Husky; cryptorchidism; BayesB analysis; genome-wide association study.

Correspondence
Max Rothschild, Department of Animal Science, Iowa State University, 2255 Kildee hall, Ames, IA 50011, USA.
Tel: +1 515 294 6202;
Fax: +1 515 294 2401;
E-mail: mfrothsc@iastate.edu

Received: 31 August 2013;
accepted: 16 October 2013

Summary

Cryptorchidism is a condition whereby one or both testes fail to descend into the scrotal sac. Here, we performed a genome-wide association study (GWAS) with both a case–control analysis using the GEMMA software accounting for population structure and a BayesB approach in the GenSel software applied to every 1 Mb window of SNPs or haplotypes. The haplotypes were constructed from a genealogical tree using the population of 204 Siberian Huskies. The BayesB analyses identified six putative genomic candidate regions on CFA6, 9, 24, 27 and X. These regions explained a high percentage of genetic variance when compared with other genomic regions. The positional candidate genes *Q9TSI5_CANFA* (matrix metalloproteinase 9 precursor) on CFA24, *ADAMTS20* (ADAM metallopeptidase with thrombospondin type 1 motif, 20) on CFA27 and *MID1IP1* (MID1 interacting protein 1) on CFAX are known to be functionally related to extracellular matrix remodelling, which might be important for gubernaculum elongation and thus interrupting normal testicular descent. Further mutation screening in these candidate regions on CFA6, 9, 24, 27 and X is needed. Next generation sequencing will help to uncover rare variants associated with cryptorchidism in this dog population.

Introduction

Testicular descent is an essential developmental stage for reproduction in male dogs. Normally, testes in the dog descend from an intra-abdominal position down into the scrotal sac on the 35th day post partum (Baumans *et al.* 1981). The environment at the scrotal sac with 2–4°C below normal body temperature is a prerequisite for normal spermatogenesis (Klonisch *et al.* 2004). Cryptorchidism is a condition of failure of testicular descent. It is considered a high risk factor for germ cell tumours in dogs (Klonisch *et al.* 2004). Moreover, bilateral cryptorchidism can result in sterility due to the thermal suppression of spermatogenesis, and it is one of the common causes of male dog infertility (Klonisch *et al.* 2004). Cryptorchidism is also regarded as one of the most common congenital human defects. Any findings discovered using the dog as a biomedical model will facilitate future cryptorchidism studies in humans.

The incidence of cryptorchidism varies in different dog breeds and populations with a reported range of 1–11% (Amann & Veeramachaneni 2007). This defect was over-represented in several pedigree dog breeds, such as German Shepherd, Boxer, Chihuahua and Siberian Husky. It has been reported that there was a high prevalence (approximately 14%) of cryptorchidism in Siberian Huskies (Zhao *et al.* 2010). The variable body location of the affected testicle(s) can cause heterogeneous phenotypes for cryptorchidism (Yates *et al.* 2003). Familial analyses showed that cryptorchidism has a genetic basis, but the mode of its inheritance remains elusive (Cox *et al.* 1978). The heritability of canine cryptorchidism is 0.22 based on a study of 2929 pure-bred boxer dogs from 414 litters when a monogenic model was applied (Nielen *et al.*

doi:10.1111/jbg.12064

2001). It is generally perceived that cryptorchidism is recessive but may involve more than one main effect gene. Other factors including epigenetic and environmental components may contribute to the occurrence of cryptorchidism (Amann & Veeramachaneni 2007).

Various genes implicated in the regulation of testicular descent have been investigated for cryptorchidism in humans, such as androgen receptor (*AR*) (Davis-Dao *et al.* 2012) and anti-Mullerian hormone (*AMH*) (Kubota *et al.* 2002). Targeted gene inactivations or mutations in rodent models have also revealed critical roles of several genes during the development of cryptorchidism. These genes in rodents included *Ar* (Kaftanovskaya *et al.* 2012), *calcitonin gene-related peptide* (Clarnette & Hutson 1999), insulin-like factor 3 *(Insl3) (*Nef & Parada 1999), relaxin/insulin-like family peptide receptor 2 (*Rxfp2*) (Gorlov *et al.* 2002), oestrogen receptor 1 (*Esr1) (*Donaldson *et al.* 1996), homeobox A10 (*Hoxa10*) (Satokata *et al.* 1995) as well as homeobox A11 (*Hoxa11)* (Hsieh-Li *et al.* 1995).

Compared with genetic evidence of cryptorchidism in humans and rodents, few studies have been performed in canine populations. Some researchers have re-examined those genes discovered in humans and mice and tried to uncover similar effects in dogs. For instance, a single nucleotide polymorphism (SNP) association study for the *ESR1* gene was performed in Miniature Dachshunds and Chihuahuas. It was reported that bilateral cryptorchid dogs might be associated with the 'GTTG' haplotype (Pathirana *et al.* 2010). Our research group has previously tested 20 candidate genes for their associations with cryptorchidism and demonstrated that collagen type II, α 1 (*COL2A1*) gene located on dog chromosome CFA27, was a potential candidate gene for cryptorchidism (Zhao *et al.* 2010).

Although studies in humans and rodents show associations of many candidate genes with cryptorchidism, the genetic basis of developing cryptorchidism in Siberian Huskies is still unknown. Therefore, in addition to our previous single SNP analysis using a candidate gene approach, we here performed a GWAS in Siberian Husky dogs employing the Illumina CanineHD BeadChip with more than 170 000 SNPs to identify genomic regions and to propose some positional candidate genes associated with variation in the incidence of cryptorchidism.

Materials and methods

Animals and phenotypes

A total of 205 male Siberian Huskies were included in this study. Among them, 43 dogs were collected from the UK, 8 from Canada and the remaining 154 were from USA. Dogs were diagnosed by palpation at 6 months of age by veterinarians or dog owners as being cryptorchids or as normal unaffected dogs. No distinction between unilateral and bilateral cryptorchidism for most of dogs was provided. In this study, there are 106 dogs with cryptorchidism and 99 diagnosed normal.

Samples and DNA preparation

Buccal samples were collected by owners from 186 dogs using sterile cytology brushes (Fisher Scientific, Hampton, NH, USA). The Canine Health Information Center (CHIC) DNA repository provided 19 male DNA samples with an average concentration of 50 ng/μl. DNA was extracted from 2 to 3 cytology brushes per dog within a single reaction of the QIAamp DNA Mini Kit (QIAGEN, Inc., Valencia, CA, USA). The DNA concentrations ranged from 10 to 50 ng/μl. Samples with DNA concentration less than 20 ng/μl were subjected to whole-genome amplification (WGA) using the GenomiPhi (GE Healthcare, Buckinghamshire, England) kit to yield high-quality DNA suitable for the Illumina Infinium canine genotyping platform.

Data generation and analyses

SNP array genotyping and quality control

DNA samples with a total DNA amount of 700–1000 ng were sent to GeneSeek, Inc. (Lincoln, NE, USA) for genotyping on the 170K CanineHD BeadChip. The SNPs with call rate ≤90%, Gentrain score ≤40% and minor allele frequency ≤0.01 were excluded from the data set as were dog samples with genotype call rate ≤80%.

Analyses of population structure

Genotypes of sampled Siberian Huskies were imported into Haploview v4.2 software (http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview). Pairwise linkage disequilibrium (LD) comparisons between markers of each chromosome were performed. Tagging SNPs were identified using the TAGGER routine implemented in Haploview.

To remove underlying group effects, we applied the model-based clustering method in STRUCTURE v2.3 software (http://pritch.bsd.uchicago.edu/structure.html) to infer population structure using SNP genotypes for all animals. The admixture model was chosen in STRUCTURE on selected tagging SNP markers, and most of the parameters were at their default values. The length of

both the burn-in and post burn-in Markov chain Monte Carlo (MCMC) iterations was 5000 with 3 runs for each K (the number of presumed population clusters). The value of K was varied from 1 to 7. To estimate the true value of K, a second order rate of changes of the likelihood function with respect to K represented as ΔK was applied. The highest value of the ΔK distribution was used as an indicator of the strength of the signal detected by STRUCTURE (Evanno *et al.* 2005). Additional analyses using principal component were also performed.

Genomewide association studies

A GWAS was carried out on all male Siberian Huskies by three approaches. The first was a case–control analysis utilizing the GEMMA software (http://home.uchicago.edu/xz7/software.html). GEMMA implements the genome-wide efficient mixed-model association algorithm and which accounts for population stratification and sample structure (Zhou & Stephens 2012). The centred relatedness matrix was first calculated from all genotypes. The association test was then performed with phenotype, genotype and the centred relatedness matrix files and fit into a univariate linear mixed model. A frequentist test called the Wald test was selected to test for the significance in GEMMA as it also considers relationship between the individuals in a data set. The second approach was a BayesB model averaging approach using the GENSEL v4R software (http://bigs.ansci.iastate.edu) with the categorical analysis option to simultaneously obtain the variance explained by SNPs in every one megabase (1 Mb) genomic window (which we refer to as the SNP model). The statistical model used for BayesB was as follows:

$$y_i = \mu + \sum_{i=1}^{k} z_{ij}\alpha_j + e_i$$

where y_i is the phenotype of animal i, μ is an overall mean, k is the total number of SNPs in the panel, z_{ij} is genotype at marker j in individual i, α_j is a random substitution effect for marker j with its own variance $\sigma^2_{\alpha_j} > 0$ (with a probability 1-π) or $\sigma^2_{\alpha_j} = 0$ (with a probability of π) and e_i is the random residual on animal i that is assumed to be independently and normally distributed N (0, σ^2_e). The parameter π was assumed to be 0.9999 so as to fit 10–20 markers (0.0001 × 170K) per iteration of the Markov chain in a mixture model for the estimation of individual SNP effects. A total of 51 050 MCMC iterations with burn-in of 1000 iterations and with an output frequency of 50 iterations were used for inference. There were 2346 unique 1 Mb windows across the dog genome based on the genotyped markers. The number of markers in each 1 Mb window varied according to genomic regions. It was assumed that under an infinitesimal model, all windows would account for the same variance of approximately 0.04%. Hence, 1 Mb windows explaining at least 0.4% of the genetic variance, which is 10-fold greater than the expected, were considered as indicative candidate regions for a gene influencing the incidence of cryptorchids.

To utilize haplotype information for identifying genes underlying complex traits, a tree-based association method that captures information about the evolutionary history (genealogy) of the genome was used in the analysis. This third approach, we refer to as the BayesB analysis with the haplotype model, was carried out to further verify regions showing associations with cryptorchidism in the BayesB analysis with the SNP model. To perform this analysis, genotypes for each individual were phased on each chromosome (based on the CanFam 2.0 assembly) using the BEAGLE v3.3 software in which a localized haplotype–cluster mode was fitted using an EM algorithm (http://faculty.washington.edu/browning/beagle/beagle.html). Subsequently, local phylogenetic trees were constructed from the phased haplotypes for each SNP marker using default values in the Blossoc software (http://www.daimi.au.dk/~mailund/Blossoc/index.html), which results in rooted binary phylogenetic trees. At the third level of each tree, there are eight possible nodes, and all haplotypes descending from each of the nodes (with the same phylogenetic lineage) were considered as a cluster. A new incidence matrix was developed to relate each of the genotyped animals to the detected haplotype clusters in addition to the original SNP genotypes. The same BayesB approach as used in the SNP model was applied to the new incidence matrix, and the posterior distributions for the genetic variance explained by each 1 Mb genomic window were obtained using the GENSEL software. The parameter π in the haplotype model was adjusted to be 0.999984 so as to fit same size of markers as in the SNP model. The windows passing the criteria (>0.40% of genetic variance) in both the SNP model and the haplotype model were retained as candidate regions in this study.

Results

There were 120 379 SNPs for GWAS analyses after quality control. One Siberian Husky dog was excluded due to its low genotype call rate and the remaining 105 affected and 99 controls were included for GWAS. A total of 47 672 tagging SNP markers were selected

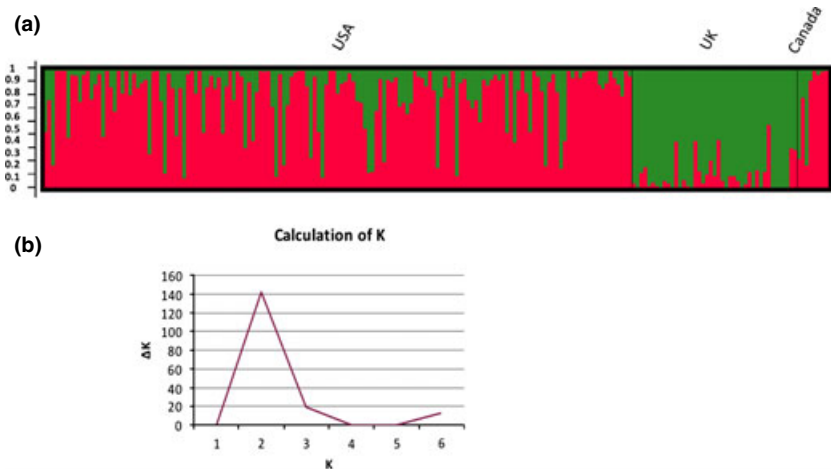

Figure 1 Detecting population structure based on the value of cluster membership estimated from STRUCTURE. (a) Summary plot of estimated Q values (cluster membership coefficients) for each individual dog collected from the USA, UK and Canada. Each individual is represented by a single vertical line broken into K coloured segments, with lengths proportional to each of the K-inferred clusters (K = 2). (b) Identification of the most probable estimate of K calculated from 47 672 tagging SNPs. Delta K was calculated for K = 1 through K = 6 under a model assuming admixture. K = 2 is the best estimate based on plot.

using Haploview (data not shown) and then used to run simulations in STRUCTURE. Under a model assuming admixture, the results recovered 2 population clusters (K = 2) according to the ΔK statistic based on the rate of change in the log probability of the data (Figure 1b). The estimated proportion of ancestry (Q value) derived from the 2 assumed population clusters varied between individual dogs (Figure 1a). Analyses using principal components produced similar results (data not shown).

The results from BayesB analyses using the SNP model identified 11 1 Mb windows explaining over 0.4% genetic variance. These 1 Mb SNP windows are on CFA3, 6, 9, 24, 27 and X. Among those 11 windows, six regions including the 31st and the 33rd Mb windows on CFAX, the 62rd Mb on CFA6, the 56th Mb on CFA9, the 36th Mb on CFA24 and the 13th Mb on CFA27 repeatedly showed a high percentage of variance (nearly or higher than 0.4%) when the windows were fitted using the haplotype model. Moreover, the results from GEMMA showed that the top SNPs with lowest raw p-values were identified within the significant windows detected from BayesB analyses. For instance, SNP BICF2P163818 in the 56th Mb on CFA9 and BICF2P514763 in the 13th Mb on CFA27, etc. (Table 1, Figure 2). A list of 53 positional characterized candidate genes were found located in these six genomic regions (Table S2). Genes called *Q9TSI5_ CANFA* (matrix metalloproteinase 9 precursor) on the 36th Mb of CFA24, *ADAMTS20* (ADAM

Table 1 The top 1 Mb windows with the highest percentage of genetic variance obtained from BayesB analyses and top SNPs detected in GEMMA

	BayesB			GEMMA	
		Genetic variance (%)			
CFA	1 Mb window (Chr_Mb)	SNP model	Haplotype model	SNP ID	p-value (Wald test)
6	**6_62****	**0.45**	**0.53**	BICF2S23120930	7.68E-04
9	9_56*	1.42	0.37	BICF2P163818	9.85E-05
24	24_36*	0.55	0.36	BICF2G630498123	2.46E-04
27	27_13*	1.14	0.34	BICF2P514763	2.11E-04
X	**X_31****	**0.71**	**0.43**	BICF2P740147	6.96E-04
X	**X_33****	**1.1**	**1.35**	BICF2S2378068	1.86E-04

**These bold windows were with genetic variance over 0.40% (10-fold than the expected value) in the 1 Mb window analyses by both SNP and haplotype models. *These windows were over 0.40% in the SNP model and nearly reached the threshold we set in the haplotype model (>0.30%). The top SNPs detected in GEMMA were coincidently located in these windows. All the above windows were considered as putative candidate regions for cryptorchidism in our study. The positional candidate genes were listed in the Table S2.

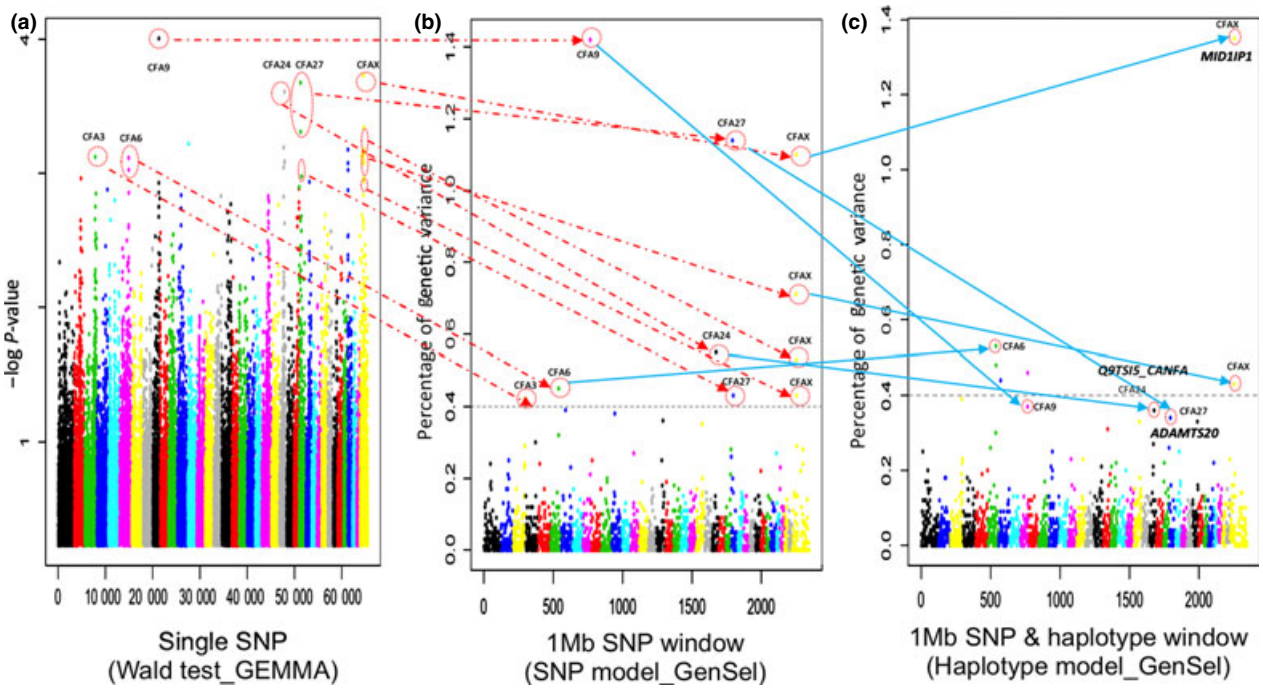

Figure 2 Whole-genome association analyses for a sample of 204 Siberian Huskies in which 105 were diagnosed with cyptorchidism. (a) Shows results of single marker analyses in GEMMA using Wald test. (b) Depicts results from the GenSel 1 Mb SNP window analyses. Each spot is a 1 Mb window containing a group of consecutive SNPs. C. Displays results using haplotypes and SNP genotypes together in GenSel. Each spot is a 1 Mb window containing a group of consecutive SNPs or haplotypes. The red breaking arrows link SNPs in GEMMA to their 1 Mb SNP windows with proportions of genetic variance greater than 0.4%. The blue solid arrows show sliding windows repeatedly identified with high variance when haplotypes were integrated into the analyses in GenSel. Six windows on CFA6, 9, 24, 27 and X were identified as putative candidate regions for cryptorchidism.

metallopeptidase with thrombospondin type 1 motif, 20) on the 13th Mb of CFA27 and *MID1IP1* (MID1 interacting protein 1) on the 33rd Mb of CFAX have been reported previously as directly or indirectly regulating extracellular matrix remodelling (Opdenakker *et al.*1998; Silver *et al.* 2008; Suzuki *et al.* 2010).

Additionally, we summarized the results from BayesB analyses for genes that had previously been shown to be associated with cryptorchidism in humans or in rodents, including *ESR1*, *NR5A1* (nuclear receptor subfamily 5, group a, member 1), *GNRHR* (gonadotropin-releasing hormone receptor), *HOXA10*, *HOXA11*, *FGFR1* (fibroblast growth factor receptor 1), *SOS1* (son of sevenless, drosophila, homolog 1), *WT-1* (Wilms tumour gene 1), *INSL3*, *AMH*, *CALCA*, *PROKR2* (prokineticin receptor 2), *LGR8*, *COL2A1*, *KAL1* (Kallmann syndrome interval gene 1) and *AR* (Table S1). The gene *COL2A1* was found 4 Mb away from the region we detected on CFA27, and the gene *NR5A1* was 5 Mb away from the window on CFA9. The BayesB analyses showed that no regions covering these genes passed the criteria set in this study, and the regions demonstrated no apparent associations with cryptorchidism. The putatively associated loci for cryptorchidism in our study do not seem to coincide with those previously reported candidate genes.

Discussion

Enormous progress has been made using GWAS to identify genetic loci associated with complex diseases and other traits of humans and animal models (Stranger *et al.* 2011). Genetic association studies can relate differences in disease status characterized as case and control groups with differences in allele frequencies at SNPs. However, undetected population substructure in an admixed population can lead to spurious allelic associations (Cardon & Palmer 2003). Diseases could be heterogeneous in different populations when the same condition is caused by different genes or different alleles at the same gene. Based on previous information, the genetic basis for cryptorchidism is likely to be heterogeneous in different populations. At least 20 genes might possibly be associated with cryptorchidism, as no single mutated gene has been found to cause more than 10% of the observed cryptorchids in the whole population (Klonisch *et al.* 2004; Amann &

Veeramachaneni 2007). Previous studies have shown that only 3–5% cryptorchid men exhibit gene aberrance for *INSL3* or *GREAT* and <16% for *AR* or *ESR1* (Amann & Veeramachaneni 2007). In all cases, it is necessary to detect the existence of an underlying population structure and then reducing the bias it might cause. Here, we used STRUCTURE to infer the presence of distinct population location clusters. The population structures were consistent for the most part with their geographic distribution (Figure 1a). The GEMMA program accounts for the population structure and provides corrected results in this study. However, the Wald test in GEMMA using a single SNP analysis is a frequentist approach that involves calculating a *P*-value for the null hypothesis of no association. The results need cautious consideration as the sample size and minor allele frequency strongly affect the results using the frequentist approach (Stephens & Balding 2009).

In this study, the Bayesian approach with SNP or haplotype windows was applied to assess evidence for an association between genomic regions containing multiple genetic variants and cryptorchidism. The Bayesian GWAS yields a posterior probability of association rather than *P*-values for the null hypothesis of no association (Stranger *et al.* 2011). Therefore, this approach can be more confidently applied than the frequentist methods when relatively small sample size exists in GWAS. In our study, we have only 205 dog samples, and while it is a relatively small population compared with many human GWAS, it was all that could be obtained from breeders and dog owners. In particular, the Bayes B method applying a threshold liability model has been confirmed to perform very well in genome-wide evaluation for a disease resistance simulation data with lower heritability (Villanueva *et al.* 2011), and this method also powerfully detected several associated QTL regions in real data with case–control and binomial disease classifications (Kizilkaya *et al.* 2013). Additionally, genomic analyses that simultaneously fit markers across the whole genome have been shown to account for structure in the population that could otherwise create spurious associations (Goddard *et al.* 2011). The Bayes B method in the GenSel program therefore is a suitable approach for this study. Hence, we primarily considered the results from Bayes B analyses as a basis for GWAS, and the results from GEMMA were treated as supplementary evidence. Most complex traits are polygenic (McCarthy *et al.* 2008). Compared with a single marker analysis strategy, multiple locus inference is a better way to identify weaker associations as small allelic effects, low minor allele frequencies as well as weak correlations with genotyped SNPs might affect the success rate of associations (McCarthy *et al.* 2008). The 1 Mb window approach in our Bayesian analyses accounts for linkage disequilibrium among adjacent SNP within the window and accumulates the effects from multiple genetic variants in the window. This method reduces the chance of spurious associations that can easily occur from single locus analysis.

Among all the possible positional candidate genes, genes *MID1IP1* on CFAX, *Q9TSI5_CANFA* on CFA24 and *ADAMTS20* on CFA27 were functionally important and may possibly be associated with cryptorchidism. *MID1IP1* encodes a protein interacting with MIDI and a TRIM/RBCC protein associated with microtubule dynamics (Berti *et al.* 2004). *MID1IP1* has also been identified associated with microtubule stabilization in Atlantic salmon (Bower & Johnston 2010). Microtubules and microfilaments determine cell shape. They are the components for cytoskeleton internally and also for the extracellular matrix externally. Extracellular matrix is rich in collagen and other macromolecules. The gubernaculum is a continuous column structure of mesenchyme connecting the testis to the scrotal sac. The swelling and migration of the male gubernaculum with relaxation of the cranial suspensory ligament in the process of testicular descent allow testes to normally move from the posterior abdominal wall to the internal inguinal ring (Hutson 1985). The extracellular matrix remodelling in/around the rodent gubernaculum has been found necessary for its migration through the inguinal fat pad with the help of matrix metalloproteinases (Churchill *et al.* 2011). Therefore, a mutated *MID1IP1* could affect microtubule stabilization and then cause abnormal extracellular matrix remodelling in/around gubernaculum. *Q9TSI5_CANFA* encodes the precursor of matrix metalloproteinase 9 (*MMP-9*). MMP-9 protein has been reported to be involved in the degradation of extracellular matrix molecules with other factors (Opdenakker *et al.* 1998). *ADAMTS20* encodes a protein belonging to the ADAMTSs (a disintegrin and metalloproteinase with thrombospondin motifs) metalloproteinase family. Mice *Adamts20* expressed in dermal mesenchymal cells has been identified to participate in melanoblast migration and survival by cleaving versican, a large extracellular matrix proteoglycan (Silver *et al.* 2008). The gubernaculum is a structure of mesenchyme. Here, we hypothesize that abnormal expression of the two matrix metalloproteinases *MMP-9* and *ADAMTS20* in dog gubernaculum might affect the extracellular matrix remodelling and/or the mesenchymal cell migration. The following disruption of gubernaculum elongation will cause abnor-

mal testicular descent. The three genes *Q9TSI5_CANFA*, *MID1IP1* and *ADAMTS20* may also interact with each other to play a role in the extracellular matrix remodelling in/around gubernaculum. To our knowledge, there is no experimental evidence for interactions between these genes.

Overall, we detected six regions on CFA6, 9, 24, 27 and X possibly associated with cryptorchidism by applying a Bayesian approach implementing window analyses and the single SNP analyses performed in GEMMA. Our results are convincing as the putative associated genomic regions explained a greater part (>0.4% for each putative region with a total of >3%) of genetic variance than those previously reported candidate genes containing genomic regions, which explained <0.15% of genetic variance. The missing genetic variance could be caused by, for example, rare variants responsible for cryptorchidism. Recent advances in next generation sequencing technologies facilitate the detection of rare variants, making it possible to uncover the roles of rare variants in complex diseases. As cryptorchidism is a complex disease and shows genetic heterogeneity, whole-genome sequencing for pairs of affected and unaffected dogs from the same family may facilitate finding rare causative mutations without needing to consider LD structure and disease heterogeneity. The obtained results let us pinpoint three positional candidate genes *Q9TSI5_CANFA, ADAMTS20* and *MID1IP1,* which are functionally related with extracellular matrix remodelling, and they may involve in gubernaculum elongation. Further experimental evidence is needed to test this hypothesis.

Acknowledgements

This work was supported by the AKC Canine Health Foundation (Grant No. 1248) and State of Iowa Funds. The help of sample collections provided by Sheila E. (Blanker) Morrissey (DVM), individual dog breeders and owners, Eddie Dziuk (Chief Operating Officer, CHIC), Caroline Kisko (secretary, Kennel Club in UK) and Dr. Nancy Bartol (Chinook Club of America) is appreciated. We also thank Ziqing Weng for help in utilizing BEAGLE program.

References

Amann R.P., Veeramachaneni D.N. (2007) Cryptorchidism in common eutherian mammals. *Reproduction*, **133**, 541–561.

Baumans V., Dijkstra G., Wensing C.J. (1981) Testicular descent in the dog. *Anat. Histol. Embryol.*, **10**, 97–110.

Berti C., Fontanella B., Ferrentino R., Meroni G. (2004) Mig12, a novel Opitz syndrome gene product partner, is expressed in the embryonic ventral midline and cooperates with Mid1 to bundle and stabilize microtubules. *BMC Cell. Biol.*, **5**, 9.

Bower N.I., Johnston I.A. (2010) Discovery and characterization of nutritionally regulated genes associated with muscle growth in Atlantic salmon. *Physiol. Genomics*, **42A**, 114–130.

Cardon L.R., Palmer L.J. (2003) Population stratification and spurious allelic association. *Lancet*, **361**, 598–604. Review.

Churchill J.A., Buraundi S., Farmer P.J., Li R., Southwell B.R., Hutson J.M., Balic A. (2011) Gubernaculum as icebreaker: do matrix metalloproteinases in rodent gubernaculum and inguinal fat pad permit testicular descent? *J. Pediatr. Surg.*, **46**, 2353–2357.

Clarnette T.D., Hutson J.M. (1999) Exogenous calcitonin gene-related peptide can change the direction of gubernacular migration in the mutant trans-scrotal rat. *J. Pediatr. Surg.*, **34**, 1208–1212.

Cox V.S., Wallace L.J., Jessen C.R. (1978) An anatomic and genetic study of canine cryptorchidism. *Teratology*, **18**, 233–240.

Davis-Dao C., Koh C.J., Hardy B.E., Chang A., Kim S.S., De Filippo R., Hwang A., Pike M.C., Carroll J.D., Coetzee G.A. (2012) Shorter androgen receptor CAG repeat lengths associated with cryptorchidism risk among Hispanic white boys. *J. Clin. Endocrinol. Metab.*, **97**, E393–E399.

Donaldson K.M., Tong S.Y., Washburn T., Lubahn D.B., Eddy E.M., Hutson J.M., Korach K.S. (1996) Morphometric study of the gubernaculum in male estrogen receptor mutant mice. *J. Androl.*, **17**, 91–95.

Evanno G., Regnaut S., Goudet J. (2005) Detection the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.*, **14**, 2611–2620.

Goddard M.E., Hayes B.J., Meuwissen T.H. (2011) Using the genomic relationship matrix to predict the accuracy of genomic selection. *J. Anim. Breed. Genet.*, **128**, 409–421.

Gorlov I.P., Kamat A., Bogatcheva N.V., Jones E., Lamb D.J., Truong A., Bishop C.E., McElreavey K., Agoulnik A.I. (2002) Mutations of the GREAT gene cause cryptorchidism. *Hum. Mol. Genet.*, **11**, 2309–2318.

Hsieh-Li H.M., Witte D.P., Weinstein M., Branford W., Li H., Small K., Potter S.S. (1995) Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility. *Development*, **121**, 1373–1385.

Hutson J.M. (1985) A biphasic model for the hormonal control of testicular descent. *Lancet*, **24**, 419–421.

Kaftanovskaya E.M., Huang Z., Barbara A.M., De Gendt K., Verhoeven G., Gorlov I.P., Agoulnik A.I. (2012) Cryptorchidism in mice with an androgen receptor

ablation in gubernaculum testis. *Mol. Endocrinol.*, **26**, 598–607.
Kizilkaya K., Tait R.G., Garrick D.J., Fernando R.L., Reecy J.M. (2013) Genome-wide association study of Infectious Bovine Keratoconjunctivitis in Angus cattle. *BMC Genet.*, **26**, 14–23.
Klonisch T., Fowler P.A., Hombach-Klonisch S. (2004) Molecular and genetic regulation of testis descent and external genitalia development. *Dev. Biol.*, **270**, 1–18.
Kubota Y., Temelcos C., Bathgate R.A., Smith K.J., Scott D., Zhao C., Hutson J.M. (2002) The role of insulin 3, testosterone, Müllerian inhibiting substance and relaxin in rat gubernacular growth. *Mol. Hum. Reprod.*, **8**, 900–905.
McCarthy M.I., Abecasis G.R., Cardon L.R., Goldstein D.B., Little J., Ioannidis J.P., Hirschhorn J.N. (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.*, **9**, 356–369.
Nef S., Parada L.F. (1999) Cryptorchidism in mice mutant for Insl3. *Nat. Genet.*, **22**, 295–299.
Nielen A.L., Janss L.L., Knol B.W. (2001) Heritability estimations for diseases, coat color, body weight, and height in a birth cohort of Boxers. *Am. J. Vet. Res.*, **62**, 1198–1206.
Opdenakker G., Fibbe W.E., Van Damme J. (1998) The molecular basis of leukocytosis. *Immun. Today*, **19**, 182–189.
Pathirana I.N., Tanaka K., Kawate N., Tsuji M., Kida K., Hatoya S., Inaba T., Tamada H. (2010) Analysis of single nucleotide polymorphisms in the 3′ region of the estrogen receptor 1 gene in normal and cryptorchid Miniature Dachshunds and Chihuahuas. *J. Reprod. Dev.*, **56**, 405–410.
Satokata I., Benson G., Maas R. (1995) Sexually dimorphic sterility phenotypes in Hoxa10-deficient mice. *Nature*, **374**, 460–463.
Silver D.L., Hou L., Somerville R., Young M.E., Apte S.S., Pavan W.J. (2008) The secreted metalloprotease ADAMTS20 is required for melanoblast survival. *PLoS Genet.*, **4**, e1000003.
Stephens M., Balding D.J. (2009) Bayesian statistical methods for genetic association studies. *Nat. Rev. Genet.*, **10**, 681–690.
Stranger B.E., Stahl E.A., Raj T. (2011) Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, **187**, 367–383.
Suzuki M., Hara Y., Takagi C., Yamamoto T.S., Ueno N. (2010) MID1 and MID2 are required for Xenopus neural tube closure through the regulation of microtubule organization. *Development*, **137**, 2329–2339.
Villanueva B., Fernández J., García-Cortés L.A., Varona L., Daetwyler H.D., Toro M.A. (2011) Accuracy of genome-wide evaluation for disease resistance in aquaculture breeding programs. *J. Anim. Sci.*, **89**, 3433–3442.
Yates D., Hayes G., Heffernan M., Beynon R. (2003) Incidence of cryptorchidism in dogs and cats. *Vet. Rec.*, **152**, 502–504.
Zhao X., Du Z.Q., Rothschild M.F. (2010) An association study of 20 candidate genes with cryptorchidism in Siberian Husky dogs. *J. Anim. Breed. Genet.*, **127**, 327–331.
Zhou X., Stephens M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.*, **44**, 821–824.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. The results of BayesB analyses on 17 previously reported candidate genes being associated with cryptorchidism

Table S2. A list of genes in the putative regions