



Genetic dissection of complex behaviour traits in German Shepherd dogs

Juliane Friedrich¹ · Erling Strandberg² · Per Arvelius³ · E. Sánchez-Molano¹ · Ricardo Pong-Wong¹ · John M. Hickey¹ · Marie J. Haskell⁴ · Pamela Wiener¹

Received: 24 May 2019 / Revised: 21 June 2019 / Accepted: 21 June 2019 / Published online: 14 October 2019
© The Author(s) 2019. This article is published with open access

Abstract

A favourable genetic structure and diversity of behavioural features highlights the potential of dogs for studying the genetic architecture of behaviour traits. However, behaviours are complex traits, which have been shown to be influenced by numerous genetic and non-genetic factors, complicating their analysis. In this study, the genetic contribution to behaviour variation in German Shepherd dogs (GSDs) was analysed using genomic approaches. GSDs were phenotyped for behaviour traits using the established Canine Behavioural Assessment and Research Questionnaire (C-BARQ). Genome-wide association study (GWAS) and regional heritability mapping (RHM) approaches were employed to identify associations between behaviour traits and genetic variants, while accounting for relevant non-genetic factors. By combining these complementary methods we endeavoured to increase the power to detect loci with small effects. Several behavioural traits exhibited moderate heritabilities, with the highest identified for Human-directed playfulness, a trait characterised by positive interactions with humans. We identified several genomic regions associated with one or more of the analysed behaviour traits. Some candidate genes located in these regions were previously linked to behavioural disorders in humans, suggesting a new context for their influence on behaviour characteristics. Overall, the results support dogs as a valuable resource to dissect the genetic architecture of behaviour traits and also highlight the value of focusing on a single breed in order to control for background genetic effects and thus avoid limitations of between-breed analyses.

Supplementary information The online version of this article (<https://doi.org/10.1038/s41437-019-0275-2>) contains supplementary material, which is available to authorised users.

- ✉ Marie J. Haskell
marie.haskell@sruc.ac.uk
- ✉ Pamela Wiener
pam.wiener@roslin.ed.ac.uk

- ¹ Division of Genetics and Genomics, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian EH25 9RG, UK
- ² Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, PO Box 7023, 750 07 Uppsala, Sweden
- ³ Swedish Armed Forces Dog Training Centre, PO Box 194, 195 24 Märsta, Sweden
- ⁴ Animal and Veterinary Sciences Group, Scotland's Rural College, Edinburgh EH25 9RG, UK

Introduction

The dog (*Canis familiaris*) is a useful animal model for identifying the genetic basis of various phenotypes (Boyko 2011; Schoenebeck and Ostrander 2014) due to its favourable genetic structure, characterised by a high linkage disequilibrium and shared haplotypes across breeds (Karlsson et al. 2007; reviewed in Hall and Wynne 2012). Behavioural traits of dogs have also been shown to have a genetic component, supported by significant within-breed genetic variance (Ilska et al. 2017), pronounced differences in behavioural characteristics between dog breeds (Mehrkam and Wynne 2014; Eken Asp et al. 2015) and Belyaev's famous "Farmed Fox" experiment in which silver foxes (close relatives of dogs) were successfully selected over several generations for increased and decreased tameness (Kukekova et al. 2012). Thus, the dog may also be a useful model for characterising the genetic architecture of behaviour and has already been used to gain insights into the genetic mechanisms underlying conditions that are also relevant in humans, such as obsessive-

compulsive disorder (Dodman et al. 2010; Tang et al. 2014). In addition to such disorders, dogs may provide unique insights into the genetic basis of complex and general behaviour characteristics, including personality traits (Hall and Wynne 2012).

There are also practical concerns for studying the genetic contribution to behaviour variation in dogs. As the first domesticated species, dogs are still employed in many roles such as herding, hunting, military and police work and serving as guide dogs, but foremost, the special social bond that developed between humans and dogs has led to the dog's popularity as a companion animal. Although dogs show tameness and strong attachment to humans in contrast to their wild ancestors, unwanted behaviours (e.g., excessive aggression, separation anxiety) still occur that affect the welfare of dogs, owners and the public (Rooney and Bradshaw 2014; Casey et al. 2014; Roth et al. 2016). Numerous studies have been performed with the aim of identifying non-genetic risk factors for the occurrence of unwanted behaviours, such as living conditions and demographic factors (Haverbeke et al. 2008; Blackwell et al. 2008; Rooney and Cowan 2011; McGreevy et al. 2013; Deldalle and Gaunet 2014; Tiira and Lohi 2015; Serpell and Duffy 2016), but few studies have considered the role of genetic factors in the management of problem behaviours. A better understanding of the genetic basis of dog behaviour may also inform breeding programmes for working dogs, e.g., guide dogs (Goddard and Beilharz 1982).

This study aims to gain general insights into the genetic architecture of behaviour variation using German Shepherd dogs (GSDs). The GSDs in this study represent unique samples of pet dogs from the United Kingdom (UK) and from a breeding programme of the Swedish Armed Forces (SAF) specifically selected for behaviour traits. By focusing on a single breed and controlling for background genetic structure that might be a consequence of analysing two populations, while also accounting for relevant environmental factors, the limitations of between-breed analyses and confounding with non-genetic effects were minimised. Moreover, different genetic approaches were applied to explore the complex nature of behaviour traits. In addition to employing a genome-wide association study (GWAS) approach based on single SNPs, a regional heritability mapping (RHM) approach was also conducted, which has been shown to perform better in the identification of multiple quantitative trait loci (QTL) with small effects (Nagamine et al. 2012). Our results highlight the complex and polygenic nature of behaviour traits and we also demonstrate that the dog is a valuable resource to study the genetic architecture of behaviour.

Materials and methods

Samples and phenotypes

Data on GSD behaviour and management was assessed using the Canine Behaviour and Research Questionnaire (C-BARQ) (Hsu and Serpell 2003) and a lifestyle survey (Friedrich et al. 2018). The C-BARQ consists of 101 questions related to training and obedience, aggression, fear and anxiety, separation-related behaviour, excitability, attachment and attention seeking and miscellaneous behaviours. The original C-BARQ was extended by 15 questions that assess the dog's playfulness (Svartberg 2005; Arvelius et al. 2014) and 21 of the miscellaneous C-BARQ questions were removed due to a lack of variability (Arvelius et al. 2014), leading to 95 final questions.

The lifestyle survey consists of questions concerning demographic factors of the dog (e.g., sex, neuter status, age), its living situation (number of children, adults and other animals living with the dog, where the dog is housed) and its current and past management (puppy socialisation, exercise and stimulation, training, activities).

Owners of registered UK GSDs that were at least 2 years-old were invited to participate in the study via email by the UK Kennel Club (KC). Participating GSDs from the UK cohort were primarily pet dogs. All GSDs from the Swedish cohort were bred within the breeding programme of the SAF. After a behaviour test at the age of 15–18 months, dogs started training for working with the SAF, Swedish Police, or other authorities or companies, and/or were selected as breeding animals, whereas others were kept as companions (Wilsson and Sinn 2012). For the Swedish cohort, owners, trainers or handlers of GSDs bred within the breeding programme of the SAF that were at least 2 years-old were invited via email or letter to participate in the study.

Behaviour data and demographic and management factors were available for 1041 GSDs from the UK and Sweden (UK = 426, Sweden = 615). To calculate the behaviour traits, a principal component analysis (PCA) was applied to the data to condense the 95 questions to a smaller number of components (described in Friedrich et al. 2018). Briefly, several procedures (Cattell's scree-test, Horn's Parallel test and the Very Simple Structure (VSS) criterion) were applied and implemented using the R package "psych" to identify the optimal number of components that capture the important information (Abdi and Williams 2010), which gave a value of 15 for all tests. The PCA was then run for 15 principal components, followed by a varimax (orthogonal) rotation (for more information see Abdi and Williams 2010). Missing values in the data set were replaced by the median value. The dogs' scores for the 15 components were

considered as quantitative behaviour traits in the subsequent analyses.

These 15 traits describe fearful, aggressive and playful behaviours in response to humans or dogs, separation anxiety, attachment and excitability, chasing, touch-sensitivity and obedience (Friedrich et al. 2018). After correcting for fixed effects (see below), the distribution of residuals for two behavioural traits, Aversion of being stepped over and Resource guarding were significantly skewed due to dogs with extreme values. A Shapiro–Wilk test of normality revealed the highest deviations from a normal distribution for the residuals of these traits and, therefore, these traits were not considered for the following analyses, leaving 13 traits for further analysis. An overview of the 13 behaviour traits (principal components) used in the subsequent analyses is given in the supplement (S1 Table).

Determination of non-genetic effects

Demographic and management factors were assessed with the lifestyle survey as described previously (Friedrich et al. 2018). Briefly, 28 factors were fitted in an initial linear model for each behaviour trait. Backward elimination was then applied to identify the model with the lowest Akaike information criterion (final model). These behaviour-specific final models were used in the subsequent analyses (S2 Table).

Genotyping and quality control

DNA was extracted for 768 dogs from saliva samples collected with Performagene PG-100 swabs (UK cohort) or blood samples (Swedish cohort) using standard protocols. The genotyping was performed using the Illumina CanineHD Whole-Genome Genotyping BeadChip featuring 172,115 SNPs. When a filter for a sample call rate of > 90% was applied, 745 dogs passed the genotyping quality control. The data set was then checked using sex and relationship information estimated from the genotype data to identify potential sampling errors and four further samples were removed. The final data set included 741 dogs (UK = 324, Sweden = 417) with sex ratios of 0.8 and 0.7 (# males: # females) for UK and Swedish dogs, respectively. SNPs were filtered in GenomeStudio software (Illumina Inc., San Diego) for call rate > 98%, reproducibility (GTS) > 0.6 and signal intensity, characterised by AB R mean (mean normalised intensity of the AB cluster) > 0.3. Using PLINK version 1.9 (Chang et al. 2015), SNPs were also filtered for minor allele frequency (MAF) > 0.05 and lack of evidence for deviations from Hardy–Weinberg equilibrium (Bonferroni-corrected P -value of $0.05 = 4.5 \times 10^{-7}$). Owing to allelic imbalance that can cause bias in association studies (discussed in Wise et al. 2013), SNPs on the X

chromosome were removed. The final set included 78,088 autosomal SNPs.

Pedigree and population structure

Although the GSDs in this study were from two different countries, there were shared pedigree links. Thus, the UK and Swedish pedigrees were merged into a joint pedigree, including both cohorts. To identify underlying population structure in the genomic data, a PCA was performed. To account for linkage disequilibrium between SNPs, a pruned SNP data set was used as input for the PCA, as recommended by PLINK version 1.9 (Chang et al. 2015). Genotype pruning on the filtered data set (78,088 SNPs) was performed using PLINK version 1.9 (Chang et al. 2015) based on the variance inflation factor, a function of the multiple correlation coefficient of a given SNP regressed on all other SNPs within a window (using default parameters: window size = 50 SNPs, the number of SNPs to shift the window at each step = 5, the variance inflation factor threshold = 2), leaving 9,180 SNPs as input for the PCA. The PCA was subsequently carried out in PLINK version 1.9 (Chang et al. 2015).

Estimation of heritability

The heritability (h^2) was estimated using pedigree and genotype data (the filtered data set of 78,088 SNPs). For the pedigree-based estimates, all GSDs with behaviour records ($n = 1041$) were used and the joint pedigree for the phenotyped dogs comprises 24,284 dogs. Heritability was estimated in ASReml (Gilmour et al. 2009) and GCTA (Yang et al. 2011) for pedigree- and genotype-based approaches, respectively, by fitting the following model:

$$y = 1\mu + Xb + Za + \epsilon \quad (1)$$

where y is a vector of behaviour traits, μ is the overall mean, b is a vector of fixed effects with X as the corresponding incidence matrix, Z is the incidence matrix for the random additive polygenic effect, a is a vector of random additive polygenic effects distributed as $MVN(0, \sigma_a^2 A)$ and $MVN(0, \sigma_g^2 G)$ for the pedigree- and genotype-based estimates, respectively, where A is the pedigree-based relationship matrix and G is the genomic relationship matrix. ϵ is a vector of residual errors distributed as $MVN(0, \sigma_e^2 I)$, where I is an identity matrix. The fixed effects include the demographic and management factors that were detected to best predict the behaviour trait (S2 Table). Dogs for which one or more fixed effects were missing were removed from the analysis, such that the number of GSDs included in the analysis varied across behaviour traits (range of 906 to 1,038 and 638 to 729 for pedigree-based and genotype-based estimations, respectively) (Table 1).

Table 1 Heritability estimates and standard deviations for behaviour traits using pedigree and genotype data

Behaviour trait ^a	Pedigree-based ^b	Genomic-based ^b
SDA: Stranger-directed aggression (12%)	0.00 ± 0.00 (1033)	0.00 ± 0.06 (729)
DDA: Dog-directed aggression (10%)	0.00 ± 0.05 (906)	0.00 ± 0.05 (638)
SDF: Stranger-directed fear (8%)	0.04 ± 0.05 (1018)	0.04 ± 0.05 (705)
Play: Human-directed playfulness (7%)	0.23 ± 0.08 (1031)	0.17 ± 0.07 (712)
EX: Excitability (7%)	0.05 ± 0.05 (1038)	0.06 ± 0.05 (725)
SA: Separation anxiety (7%)	0.00 ± 0.00 (1010)	0.00 ± 0.05 (716)
LO: Lack of obedience (7%)	0.00 ± 0.00 (1011)	0.00 ± 0.06 (711)
SDI: Stranger-directed interest (6%)	0.10 ± 0.06 (985)	0.01 ± 0.05 (687)
AS: Attachment/ Attention seeking (6%)	0.00 ± 0.00 (1003)	0.02 ± 0.05 (706)
CH: Chasing (6%)	0.09 ± 0.06 (966)	0.13 ± 0.06 (659)
NSF: Non-social fear (6%)	0.12 ± 0.06 (1025)	0.16 ± 0.06 (727)
DDF: Dog-directed fear (5%)	0.01 ± 0.04 (1001)	0.00 ± 0.04 (698)
TS: Touch-sensitivity (4%)	0.02 ± 0.04 (966)	0.00 ± 0.04 (672)

Values highlighted in bold are significant according to a log-likelihood ratio test (for pedigree-based estimations) or for $P < 0.05$ (for genomic-based estimations)

^aBehaviour traits are listed in descending order for the variance explained by the PCA used to define the traits (given in parentheses; see Supplement 1)

^bHeritability estimates with the number of dogs in the analysis (given in parentheses)

The significance of pedigree-based h^2 was tested using a log-likelihood ratio test (LRT) in ASReml (Gilmour et al. 2009), comparing the log-likelihood ratio statistic to a χ^2 (d.f. = 1) for $P < 0.05$. The significance of genotype-based estimates was defined by P -values < 0.05 from the LRT within the genome-based restricted maximum likelihood (GREML) analysis performed in GCTA (Yang et al. 2011).

Genome-wide association study (GWAS)

A GWAS was performed on the filtered data set of 78,088 SNPs to identify associations between SNPs and behaviour traits based on an additive model. To account for population structure, models with different combinations of factors (cohort as fixed effect, genotype-derived principal components 1 and 2 as covariates, genomic relationship matrix as random effect) were evaluated. Fitting only the cohort and the relationship matrix performed best, as assessed by the genomic inflation factor (λ) (i.e., closest to 1.0). The following linear model was fitted in GEMMA (Zhou and Stephens 2012):

$$y = 1\mu + Xb + c\beta + Za + \varepsilon \quad (2)$$

where y is a vector of behaviour traits, μ is the overall mean, b is a vector of fixed effects with X as the corresponding incidence matrix, c is a vector of marker genotypes (alleles coded as 0/1) with β as the vector of regression coefficients of the phenotype on the marker genotypes, Z is the incidence matrix for the random additive polygenic effect, a is a vector of random additive polygenic effects with $MVN(0, \sigma_a^2 G)$, where G is the genomic relationship matrix, and ε is a vector of residual errors with $MVN(0, \sigma_e^2 I)$, where

I is an identity matrix. The fixed effects comprise the demographic and management factors obtained in the individual final models (S2 Table).

A conservative Bonferroni correction was applied to determine genome-wide significance ($P < \frac{0.05}{78,088}$; 6.4E-07) and suggestive (allowing one false-positive per genome scan: $P < \frac{1}{78,088}$; 1.3E-05) (Riggio et al. 2013) thresholds that account for the multiple testing resulting from the large number of markers but not for multiple behaviour traits.

Regional heritability mapping (RHM)

Genomic regions were also tested for association with behaviour traits. This was carried out by scanning windows across the whole-genome using RHM, performed in REACTA (Grey et al. 2012). This approach used the model described by Nagamine et al. (2012) where two genetic effects are fitted: the first representing the overall genetic effects (modelled with an overall genomic relationship matrix calculated using all SNPs across the genome) and the second genetic effect representing the effect associated with the specific region of the genome being tested (modelled with a regional genomic relationship matrix calculated using only SNPs from this region). The SNPs used for the regional relationship matrix were excluded from the overall genomic relationship matrix (Cebamanos et al. 2014). REACTA (Grey et al. 2012) uses a sliding-window approach and we used a fixed window size of 50 SNPs with overlaps of 25 SNPs. The window size of 50 SNPs was chosen as a compromise between power to detect associations and computational demands (Uemoto et al. 2013).

Using these parameters resulted in 3124 regions under analysis; to correct for multiple testing, a Bonferroni correction was applied to genome-wide significance ($P < \frac{0.05}{3124}$; 1.6E-05) and suggestive ($P < \frac{1}{3124}$; 3.1E-04) thresholds.

Analysis of candidate genes and regions

The coordinates of identified SNPs and regions were mapped to the CanFam3.1 assembly to identify (I) genes harbouring or near identified SNPs (GWAS) and (II) genes located within identified regions (RHM). Regarding (I): to determine the size of the region around identified SNPs that should be scanned for candidate genes, the squared correlation (r^2) between all pairs of SNPs within 10 Mb were calculated across the genome using PLINK version 1.9 (Chang et al. 2015). The average r^2 was calculated for bins of increasing distance between SNPs to identify the distance around SNPs at which average r^2 drops below 0.5. The longest bin for which average $r^2 > 0.5$ was 200 kb and thus this distance was chosen as the region around associated SNPs to be investigated. Regarding (II), the GWAS results, $-\log_{10}(P)$, were plotted within the regions identified by RHM to identify positional candidate genes. The pairwise r^2 was calculated between all SNPs in the region and the SNP with highest $-\log(P)$ value to describe the pattern of linkage for the region, using PLINK version 1.9 (Chang et al. 2015) as described above. The regional associations plots were created using an R script modified from that of Saxena et al. (2007).

All genes within the regions described above (I and II) were submitted to Enrichr (Chen et al. 2013; Kuleshov et al. 2016) to identify enriched biological processes.

Results

Population structure

We explored the underlying population structure in the two GSD cohorts by applying a PCA to the genomic data. The variance in the genomic data explained by the first three principal components was 2.18%, 1.68% and 1.22%, respectively, and 66.96% of the variance was explained by all components with eigenvalue > 1 . Plotting the first two components of the PCA (Supplementary Fig. S3) shows population structure due to cohort by a clear separation of UK and Swedish dogs based on the first principal component. However, some GSDs overlapped between the cohorts, showing shared ancestry. In contrast to the cohort effect, there were no distinct patterns observable for eigenvectors PC1 and PC2 when considering the GSDs according to their function or coat colour.

Heritabilities

Heritability estimates for the 13 behaviour traits were calculated using pedigree and genomic data. Moderate and significant h^2 were found for Human-directed playfulness and Non-social fear using pedigree and genomic approaches, while Stranger-directed interest was only significant for pedigree-based estimates and Chasing only for genomic estimates (Table 1). The highest h^2 were calculated for Human-directed playfulness using pedigree data (0.23 ± 0.08) and for Non-social fear using genotype data (0.16 ± 0.06). Non-significant heritabilities were estimated for Stranger-directed fear, Excitability, Attachment/ Attention seeking, Dog-directed fear and Touch-sensitivity using estimates from pedigree and genomic data.

Association mapping

Genome-wide association studies (GWAS) and a regional heritability mapping (RHM) were performed as complementary approaches to identify associations between genetic markers and the 13 behaviour traits (Fig. 1). The average genomic inflation for GWAS across the 13 behaviour traits was 0.99 (ranging from 0.89 to 1.06), showing that population stratification was adequately controlled (Supplementary Fig. S4). In the GWAS, a total of 15 SNPs were found with a suggestive association to one of the analysed behaviour traits and two of these also showed a genome-wide significant association ($P < 6.4E-07$) (Table 2).

The identified SNPs were distributed over seven of the 38 canine autosomes, with the largest numbers on CFA33 (5) for Attachment/Attention seeking, CFA31 (3) for Dog-directed fear and CFA14 (3) for Stranger-directed interest. The genome-wide associations were found for Attachment/ Attention seeking (two adjacent SNPs on CFA33). The greatest number of suggestive SNPs were found for Attachment/ Attention seeking (6), Stranger-directed interest (3) and Dog-directed fear (3).

The RHM analysis was performed by testing for associations between 50-SNP sliding windows across the genome (with a 25-SNP overlap between consecutive windows) (Fig. 1). Scanning the genome for regions associated with the 13 behaviour traits based on the suggestive threshold, we identified 16 regions associated with at least one of the behaviour traits (Table 3). One region on CFA33 associated with Attachment/Attention seeking showed genome-wide significance and also harbours the only SNPs with genome-wide significance in the GWAS. The average size of the identified regions was 1.31 Mb (range: 0.89–2.63 Mb).

Most of the SNPs identified by the GWAS overlapped with regions identified by the RHM (Tables 2 and 3 and Fig. 1), only the SNPs found on CFA10 and CFA17 for

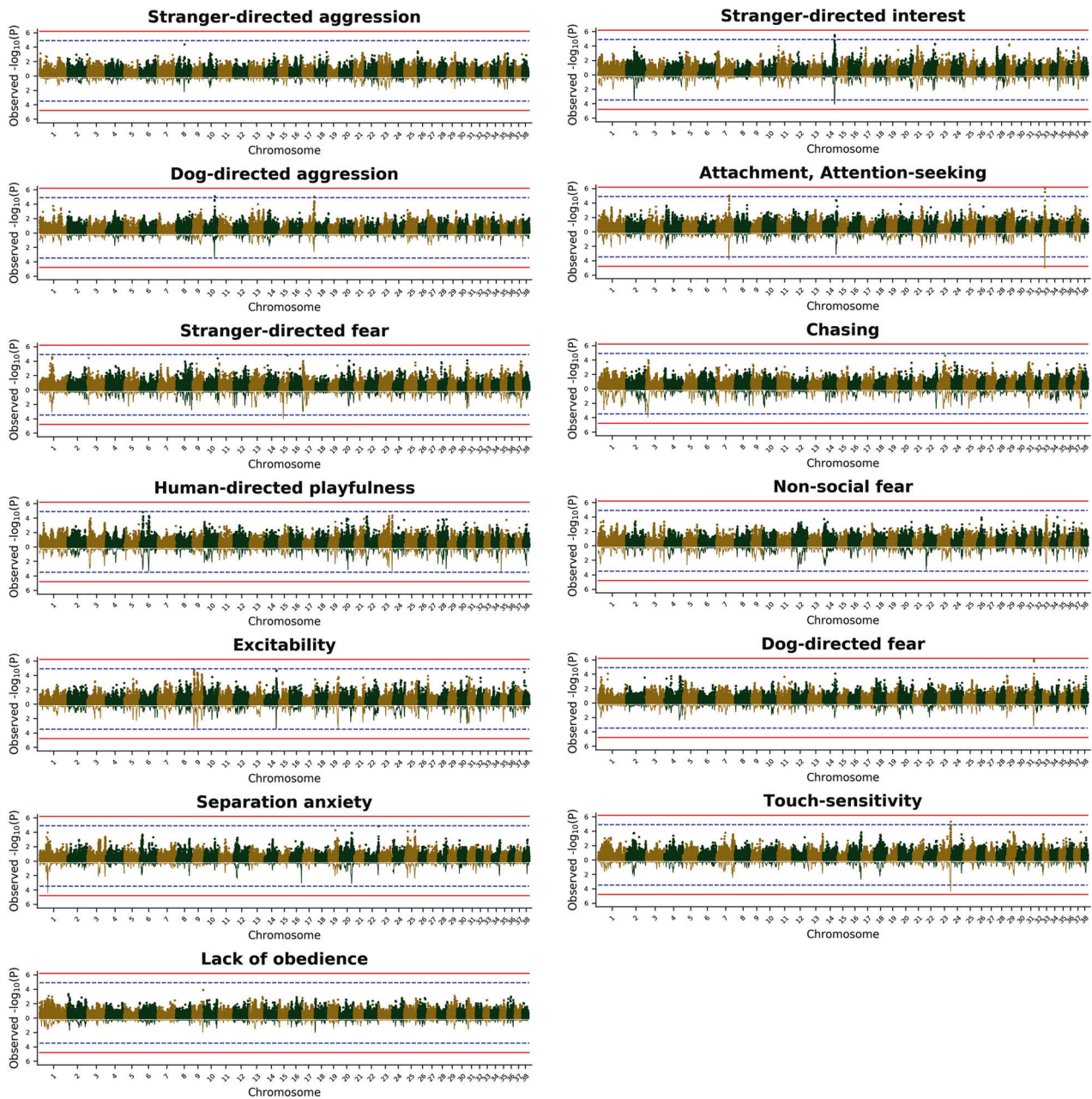


Fig. 1 Joint Manhattan plots for GWAS and RHM analyses for the 13 analysed behaviour traits. Negative log P -values for each SNP and region were plotted according to their chromosomal position for the

Dog-directed aggression and on CFA31 for Dog-directed fear were exclusive to the GWAS approach. Exclusive peaks were also found with the RHM approach, for example on CFA1 for Separation-anxiety, on CFA3 for Chasing, and on CFA19 for Excitability.

Candidate genes and regions

According to the annotation of CanFam3.1, four of the SNPs identified by the GWAS were located within three genes

GWAS (upper plot) and the RHM (lower plot) for each behaviour trait. The red line indicates the genome-wide significance threshold and the blue dotted line indicates the suggestive threshold

(*ARNT*, *PLCH1* and *BRWD1*) and 30 genes were located within 200 kb of suggestive or genome-wide significant SNPs (Table 2). The two SNPs on CFA33 with genome-wide significance for Attachment/Attention seeking were located approximately 63 kb downstream of an unannotated protein-coding gene (*ENSCAFG00000009706*). Gene ontology analysis of the 30 genes revealed that the top enriched biological processes were “polyphosphate metabolic process” (GO: 0006797; adjusted P -value = 0.009), “negative regulation of axon regeneration” (GO: 0048681;

Table 2 Significant and suggestive results for the genome-wide association study

Trait	Chr	Pos (bp)	SNP name	β	<i>p</i> -value	Gene(s)
AS	7	62043815	BICF2G630564528	0.27 ± 0.06	9.79E-06	<i>AQP4</i> ; <i>KCTD1</i>
DDA	10 ^a	52699559	BICF2G630490170	0.25 ± 0.06	8.06E-06	<i>ENSCAFG00000002716</i>
SDI	14	49407681	BICF2P1296430	−0.23 ± 0.05	2.78E-06	
SDI	14	49546253	BICF2S22928246	−0.23 ± 0.05	4.32E-06	
SDI	14	49864037	BICF2P325193	−0.22 ± 0.05	4.73E-06	
DDA	17 ^a	59982961	BICF2P588067	−0.45 ± 0.10	9.10E-06	<i>ANXA9</i> ; <i>ARNT</i> ; <i>BNIP1</i> ; <i>C17H1orf56</i> ; <i>CDC42SE1</i> ; <i>CERS2</i> ; <i>CTSK</i> ; <i>CTSS</i> ; <i>ENSA</i> ; <i>GABPB2</i> ; <i>GOLPH3L</i> ; <i>HORMAD1</i> ; <i>MINDY1</i> ; <i>MLLT11</i> ; <i>PRUNE1</i> ; <i>SETDB1</i>
TS	23	49477874	BICF2G630364231	0.30 ± 0.07	4.82E-06	<i>PLCHI</i> ; <i>C3orf33</i> ; <i>SLC33A1</i> ; <i>GMPS</i>
DDF	31 ^a	34182127	BICF2P766705	−0.34 ± 0.07	9.17E-07	<i>B3GALT5</i> ; <i>BRWD1</i> ; <i>HMGNI</i> ; <i>ENSCAFG00000010048</i>
DDF	31 ^a	34203899	BICF2P402445	−0.34 ± 0.07	9.17E-07	<i>B3GALT5</i> ; <i>BRWD1</i> ; <i>HMGNI</i> ; <i>ENSCAFG00000010048</i>
DDF	31 ^a	34242284	BICF2P544489	−0.32 ± 0.07	1.66E-06	<i>B3GALT5</i> ; <i>BRWD1</i> ; <i>HMGNI</i> ; <i>ENSCAFG00000010048</i>
AS	33	9217774	BICF2G630248670	0.35 ± 0.07	1.02E-06	<i>ENSCAFG00000009682</i>
AS	33	9686098	BICF2G630248946	0.41 ± 0.08	1.64E-07 ^b	<i>ENSCAFG00000009697</i> ; <i>ENSCAFG00000009706</i>
AS	33	9691453	BICF2G630248954	0.41 ± 0.08	1.64E-07 ^b	<i>ENSCAFG00000009697</i> ; <i>ENSCAFG00000009706</i>
AS	33	9705043	BICF2G630248964	0.38 ± 0.08	2.94E-06	<i>ENSCAFG00000009706</i>
AS	33	9705526	BICF2G630248967	0.38 ± 0.08	2.94E-06	<i>ENSCAFG00000009706</i>

Trait: Behaviour trait

Chr: chromosome number

SNP name: SNP ID

β : effect size and standard deviation

SE standard error or beta estimate

Gene(s): Genes located +/− 200 kb around SNP (genes harbouring the SNP are highlighted in bold)

Coordinates, statistics of the REML analysis and positional candidate genes are given for all SNPs that exceeded the suggestive or genome-wide significance threshold

^aLocus was exclusively identified by the genome-wide association study and not by the regional heritability mapping for the same trait (all other significant SNPs were located within significant regions)

^bGenome-wide significant association

adjusted *P*-value = 0.12) and “regulation of hormone biosynthetic process” (GO: 0046885; adjusted *P*-value = 0.12).

To further investigate regions identified by the RHM analysis, $-\log(P)$ values obtained from the GWAS, gene annotations and local linkage disequilibrium patterns were plotted for these regions to pinpoint the most likely location of positional candidate genes (Supplementary Fig. S5). Overlapping regions, due to the sliding-window approach of the RHM analysis, were combined. There were 60 genes located in these regions (Table 3); of these, several functional candidate genes (*LRRN3*, *KCNAB1* and *BRWD1*) were also located near (Supplementary Fig. S5) or at (Table 2) SNPs identified by GWAS. Two other functional candidate genes (*HIVEP2* and *AIG1*) were located in identified regions but the $-\log(P)$ values for nearby SNPs obtained in the GWAS did not exceed the suggestive threshold (Supplementary Fig. S5). The region on CFA33 with genome-wide significance for Attachment/Attention seeking comprised three unannotated protein-coding genes (*ENSCAFG00000009682*, *ENSCAFG00000009697* and *ENSCAFG00000009706*).

According to the gene ontology analysis, the GO biological processes significantly enriched by genes located in identified regions (Table 3) were “histidine catabolic process” (GO: 0006548; adjusted *P*-value = 0.013), “histidine metabolic process” (GO: 0006547; adjusted *P*-value = 0.013) and “imidazole-containing compound catabolic process” (GO: 0052805; adjusted *P*-value = 0.013).

Discussion

Dogs express diverse behaviour phenotypes, some of which appear to be related to traits of other species (including humans), making them useful models for general insights into the genetic architecture of behaviour. However, behaviours are complex traits, which have been shown to be influenced by numerous non-genetic (environmental) factors and genetic variants of low to moderate effect (Flint 2003), which complicates their analysis and the identification of underlying genes and mechanisms. In this study, we

Table 3 Significant and suggestive results for the regional heritability mapping

Trait	Chr	Start (Mb)	Stop (Mb)	h^2	LRT	P-value	Gene(s)
SA ^a	1	33.9	34.8	0.19 ^b ± 0.09	15.5	4.05E-05	<i>ADGRG6</i> ; <i>AIG1</i> ; <i>HIVEP2</i>
CH ^a	3	8.72	10.3	0.03 ± 0.03	13.5	1.19E-04	<i>FAMI74A</i> ; <i>SLCO4C1</i> ; <i>SLCO6A1</i> ; <i>ST8SIA4</i>
CH ^a	3	9.68	11.0	0.03 ± 0.02	11.8	3.03E-04	<i>FAMI74A</i> ; <i>ST8SIA4</i>
AS ^a	7	61.4	62.5	0.03 ± 0.02	12.3	2.31E-04	<i>AQP4</i> ; <i>CHST9</i> ; <i>KCTDI</i> ; <i>PSMA8</i> ; <i>SSI8</i> ; <i>TAF4B</i>
AS	7	62.0	63.0	0.04 ± 0.02	13.3	1.33E-04	<i>KCTDI</i> ; <i>PSMA8</i> ; <i>SSI8</i> ; <i>TAF4B</i>
SDI	14	49.1	50.1	0.02 ± 0.01	13.6	1.16E-04	
SDI	14	49.6	50.6	0.05 ± 0.03	14.3	7.81E-05	
SDI ^a	14	50.1	51.3	0.05 ± 0.03	13.3	1.35E-04	<i>ENSCAFG00000003294</i> ; <i>ENSCAFG000000003299</i> ; <i>LRRN3</i>
SDF ^a	15	35.5	37.7	0.18 ^b ± 0.10	12.8	1.69E-04	<i>AMDHDI</i> ; <i>CCDC38</i> ; <i>CDK17</i> ; <i>ELK3</i> ; <i>ENSCAFG000000006395</i> ; <i>ENSCAFG000000006472</i> ; <i>ENSCAFG000000006487</i> ; <i>ENSCAFG00000006494</i> ; <i>HAL</i> ; <i>LTA4H</i> ; <i>NEDD1</i> ; <i>NTN4</i>
SDF ^a	15	36.9	39.5	0.13 ^b ± 0.08	13.9	9.44E-05	<i>ACTR6</i> ; <i>ANKS1B</i> ; <i>APAF1</i> ; <i>ENSCAFG000000006545</i> ; <i>ENSCAFG000000006667</i> ; <i>ENSCAFG000000006682</i> ; <i>ENSCAFG000000023440</i> ; <i>ENSCAFG00000006734</i> ; <i>IKBIP</i> ; <i>SCYL2</i> ; <i>SLC17A8</i> ; <i>TMPO</i> ; <i>UHRF1BP1L</i>
Ex ^a	19	48.4	49.2	0.04 ± 0.03	11.7	3.06E-04	
TS	23	48.6	50.0	0.04 ± 0.03	13.4	1.26E-04	<i>C3orf33</i> ; <i>GMPS</i> ; <i>KCNABI</i> ; <i>MME</i> ; <i>PLCHI</i> ; <i>SLC33A1</i>
TS	23	49.3	51.0	0.04 ± 0.03	15.4	4.24E-05	<i>C3orf33</i> ; <i>CCNLI</i> ; <i>ENSCAFG00000029908</i> ; <i>ENSCAFG00000008882</i> ; <i>ENSCAFG00000032528</i> ; <i>GMPS</i> ; <i>KCNABI</i> ; <i>PLCHI</i> ; <i>PTX3</i> ; <i>SLC33A1</i> ; <i>SSR3</i> ; <i>TIPARP</i> ; <i>VEPH1</i>
AS	33	8.18	9.57	0.05 ± 0.03	12.5	2.09E-04	<i>ZPLDI</i>
AS	33	8.85	10.1	0.07 ± 0.04	18.6	7.90E-06 ^c	<i>ENSCAFG00000009682</i> ; <i>ENSCAFG00000009697</i> ; <i>ENSCAFG00000009706</i>
AS	33	9.60	10.6	0.07 ± 0.04	13.5	1.20E-04	<i>ENSCAFG00000009706</i> ; <i>ENSCAFG00000025525</i>

Trait: Behaviour trait

Chr: chromosome number

Start (Mb) and Stop (Mb): beginning and end of the region

 h^2 : estimate of heritability and standard deviation

Coordinates, statistics of the association analysis, regional heritabilities and positional candidate genes are given for all genomic regions that exceeded the suggestive or genome-wide significance threshold. Owing to the sliding-window approach used in the analysis, the regions comprise 50 SNPs and can overlap with adjacent regions by 25 SNPs

SE standard error of heritability, LRT log-likelihood ratio

^aLocus was exclusively identified by the regional heritability mapping and harboured no significant SNP identified by the GWAS for the same trait^bOverestimation (regional h^2 is inconsistent with the corresponding genome-wide h^2)^cGenome-wide significant association

analysed the influence of genetic factors on behaviour traits of German Shepherd dogs using multiple genomic approaches, while accounting for various non-genetic factors, with the aims of characterising the general genetic architecture of behaviour and identifying candidate genes.

The genetic contribution to behaviour variation

The heritabilities estimated for the 13 behaviour traits using pedigree and genomic approaches ranged from 0 to 0.23. These measures for h^2 are within the range of most previously observed values in dogs (Saetre et al. 2006; Arveilius et al. 2014; Ilska et al. 2017), while a few studies reported higher h^2 for similar behaviour traits (Ruefenacht et al. 2002; van der Waaij et al. 2008). Discrepancies between observed h^2 for dog behaviour traits across studies can be explained by the different behaviour phenotypes used, e.g., whether the behaviour was subjectively scored or actually measured and whether the behaviour was recorded in everyday life or in test situations, and also by differences between breeds (due to different population histories).

From other species it is known that specific behaviour patterns contributing to the fitness of an individual, such as courtship or feeding, are under stronger genetic control than behaviours with apparently less evolutionary relevance like personality traits (York 2018). In this study, behaviour traits with substantial h^2 were Human-directed playfulness, Non-social fear, Stranger-directed interest and Chasing. The observation of the highest h^2 across traits for Human-directed playfulness has been also made in a genetic study of 14 different dog breeds (Asp et al. 2014). While many other studies on the genetic background of dog behaviour focused on human-directed aggression (Liinamo et al. 2007; Våge et al. 2010; Zapata et al. 2016), we included traits of playful interactions in our analysis since playfulness in regard to humans has been shown to explain a large proportion of the variance between individuals in the analysis of multiple dog breeds (Svartberg 2005). In particular, Human-directed playfulness and Stranger-directed interest describe boldness and attachment to humans and our results indicate that these behaviour characteristics might be directly targeted by selection for tameness and human-attachment in dogs. Specifically regarding GSDs, although the SAF do not use C-BARQ for their selection programme, a previous study showed significant associations between success in a temperament test assessing dogs for further training and C-BARQ-measured traits of young dogs related to Lack of obedience, Stranger-directed fear, Non-social fear, Dog-directed fear and Touch-sensitivity (Foyer et al. 2014), suggesting that these traits have been selected against in the Swedish cohort. We do not have similar information for the UK cohort as these dogs are primarily pets and not part of a breeding programme, however, it is

possible that selection criteria over recent years have been based more on cosmetic traits as the breed has moved from a working dog to pet (O'Neill et al. 2017).

Using genome-wide association and regional heritability mapping, we identified 15 SNPs and 16 regions, respectively, which showed suggestive association with one of the analysed behaviour traits. These SNPs and regions were distributed over 11 chromosomes. Several regions were identified by both GWAS and RHM.

Comparing genomic regions identified in the current study to the results from other single-breed studies, we found that the SNP for Attachment/Attention seeking on CFA7 is located in a region of ~1 Mb flanked by two loci associated with obsessive-compulsive disorder in Doberman Pinschers (Tang et al. 2014). In contrast, the suggestive SNPs identified for behaviour traits in Labrador Retrievers by Ilska et al. (2017) do not overlap with candidate regions found in the current study. Furthermore, none of the genetic regions mapped to aggression and fear across multiple dog breeds in a study by Zapata et al. (2016) overlapped with genetic regions found in the current study. Ostrander et al. (2017) reviewed the identified loci for behaviour traits across dog breeds by Zapata et al. (2016) and found that many of these loci were previously linked to body size, suggesting that behaviour may have been confounded with physical characteristics in between-breed analyses or an association between behaviour and some morphological traits. In the silver fox experiment described above, changes in behaviour were also accompanied by physiological and morphological changes (Trut 1999) and other studies have shown an association between behaviour and body traits across breeds (McGreevy et al. 2013), suggesting an genetic interplay between these traits. These observations might also indicate that GWAS across dog breeds are more sensitive for morphological differences than for variation in behaviour, which highlights the importance of single-breed analyses in the dissection of the genetic background of behaviour. In contrast to the Zapata et al. (2016) study, candidate regions identified in the current study do not overlap with known genetic regions associated with body size (based on the largest study to date, Hayward et al. 2016).

However, our results also suggest that QTL for dog behaviour may be breed-specific as indicated by the lack of QTL that overlap those found in other studies. It is likely that across breeds, different behaviour-oriented breeding practices have led to different alleles selected to moderate frequencies, leading to breed-specific QTL.

Candidate genes related to behaviour traits

In this study, we combined two complementary approaches (GWAS and RHM) with the aim of detecting novel

candidate genes for behaviour and further evaluating genes previously linked to behaviour.

The only SNPs and region with genome-wide significance for the behaviour trait Attachment/Attention seeking point to a region on CFA33 that contains several unannotated protein-coding genes, including *ENSCAFG0000009706*. According to the iDOG database (Tang et al. 2019), *ENSCAFG0000009706* is a protein-coding gene with molecular functions related to RNA binding and the structural constitution of the ribosome (GO: 0003723 and 0003735). However, this gene has not yet been described in other canine association mapping studies.

Many of the other positional candidate genes have been previously linked to behaviour characteristics and disorders or to neuronal development, especially in regards to humans. The aquaporin-4 (*AQP4*) gene identified by both GWAS and RHM for Attachment/Attention seeking is one of the most abundant molecules in the brain, with many physiological functions (reviewed in Nagelhus and Ottersen 2013). In a study on gene expression changes in the brains of dogs and wolves, *AQP4* showed a significant fourfold higher gene expression in dog than in wolf, indicating that it may have played a role in domestication (Saetre et al. 2004). Our results provide further evidence for the role of this gene regarding attachment to humans.

RHM identified several regions that were not identified by the GWAS and contain genes that have previously been linked to behaviour. The region at ~34 Mb on CFA1, associated with Separation anxiety, includes *HIVEP2* and *AIG2*, which have been previously identified as positional candidate genes in a GWAS on affiliative social behavior in humans (Knoll et al. 2018). The region at 50–52 Mb on CFA14, associated with Stranger-directed interest, includes *LRRN3*, a strong risk gene for autism in humans (Hutcheson et al. 2004). In addition, the region at ~49–51 Mb on CFA23, associated with Touch-sensitivity (a behaviour trait that is characterised by fearful or aggressive responses to grooming or bathing), contains another promising functional candidate gene, *KCNAB1*. Two SNPs with low but not quite suggestive *P*-values in the GWAS were also located within the *KCNAB1* gene, which encodes the voltage-gated potassium channel subunit beta-1. Interestingly, mouse knockouts at the *KCNQ* gene, which encodes another voltage-gated potassium channel, showed an increased sensitivity of mechanoreceptors in the skin (Schütze et al. 2016). It is possible that variation in *KCNAB1* could have a similar effect and thus this might influence touch-sensitivity in dogs.

The GO analysis for genes identified by the RHM revealed an enrichment of catabolic and metabolic histidine processes due to the genes *AMDHD1* and *HAL* (the region harbouring these two genes was associated with Stranger-directed fear). Histidine is a precursor of the

neurotransmitter histamine and it has been shown that the histaminergic system affects the central nervous system and thus also alters behaviours, e.g., by affecting fear memory (reviewed in Passani et al. 2007).

Other genes were identified only by the GWAS, including *BRWD1* (CFA31), *B3GALT5* (CFA31) and *ARNT* (CFA17). Two SNPs associated with Dog-directed fear are located within *BRWD1*. In human GWAS studies, this gene has been associated with cognitive function (Davies et al. 2018), intelligence (Savage et al. 2018) and temperament in individuals with a bipolar disorder (Greenwood et al. 2012). In close proximity to these SNPs lies *B3GALT5*, which has been linked to suicide attempts (Perlis et al. 2010) and obsessive-compulsive symptoms (den Braber et al. 2016). Finally, a SNP on CFA17 associated with Stranger-directed interest is located within the *ARNT* gene. Variation within *ARNT* has been linked to the severity of autism in humans (Fujisawa et al. 2016).

Limitations and implications for further studies

The limited number of genome-wide significant associations found in this study indicates the challenges in the genetic dissection of complex traits like behaviour, which derive from the small effects of genetic variants on phenotypic variation, substantial environmental effects and difficulties in defining clear phenotypes. Although ours is one of the largest genomic studies of dog behaviour so far, it has been shown in human studies that much larger sample sizes are required for robust genetic dissection of complex traits, e.g., height (Visscher et al. 2014). The use of C-BARQ, a standardised owner-derived questionnaire, to measure behaviour phenotypes, which has been successfully applied in many studies and records a range of behaviours in everyday situations, opens the possibility of meta-analysis across studies and thus ultimately achieving a larger sample size. However, a limitation of using questionnaire-based phenotypes is that the recorded traits are influenced by the subjectivity of the participants, which might be even more pronounced when participants originate from different countries and thus show cultural differences as in this study. While we attempted to correct for this in the statistical analysis, we may not have been completely successful.

Conclusions

Understanding the genetics of dog behaviour and the interaction with non-genetic factors can give general insights into animal and human behaviour and is relevant for animal welfare, e.g., to identify risk factors for problem behaviours. Our results support the hypothesis that

behaviours are complex traits, influenced by multiple genetic and non-genetic factors, emphasising the need for large datasets incorporating both genetic and non-genetic information in future studies of dog behaviour. Furthermore, it is important to reach a consensus on the non-genetic factors with greatest effects on these traits in order to standardise analyses.

If these requirements are met, dogs can provide a valuable resource for studying the genetics of behaviour characteristics, especially in terms of intra- and inter-species social interactions. In this study, genomic regions and SNPs associated with behaviour traits suggested a number of candidate genes that were previously described for psychological disorders in humans, indicating a potential new context for these genes in the general expression of behaviour variation. By analysing a single dog breed, we were able to highlight candidate genes for behaviour that are less likely to be confounded with morphological variation compared to between-breed analyses. However, further studies with larger sample sizes are required to identify and confirm the identified associations and candidate genes and, where associations are confirmed, subsequent functional analyses will be needed to progress in understanding how these genes influence expression of behaviour.

Data archiving

Data for the UK dogs is available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.493rk16>.

Acknowledgements We want to thank all owners of German Shepherd dogs participating in this study for their time and effort to answer the questionnaires and send saliva samples for genotyping. Thanks are also extended to the Kennel Club, the British Association for German Shepherd Dogs, and the German Shepherd Dog Breed Council of Great Britain for assistance in participant recruitment for the UK cohort. Thanks to Zita Polgar, Carol-Anne Duthie and Joanna Warner for assistance in contacting dog owners. We would also like to thank the SAF Dog Training Centre, in particular Lisa Rutström, for recruiting participants for the Swedish cohort, and Susanne Gustafsson and Gabriela Bottani Claros (Swedish University of Agricultural Sciences) for providing DNA samples. We thank Helen Brown for statistical advice and Dr. James Serpell (University of Pennsylvania, USA) for permission to use C-BARQ. Primary funding was provided by the Dogs Trust (UK); further funding was provided by BBSRC Institute Strategic Programme Grants (to the Roslin Institute) and RESAS, Scottish Government (to SRUC).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdi H, Williams LJ (2010) Principal component analysis. *Wiley Interdiscip Rev Comput Stat* 2:433–459
- Arvelius P, Asp HE, Fikse WF, Strandberg E, Nilsson K (2014) Genetic analysis of a temperament test as a tool to select against everyday life fearfulness in Rough Collie. *J Anim Sci* 92:4843–4855
- Arvelius P, Strandberg E, Fikse WF (2014) The Swedish Armed Forces temperament test gives information on genetic differences among dogs. *J Vet Behav* 9:281–289
- Asp HE, Arvelius P, Fikse WF, Nilsson K, Strandberg E (2014). Genetics of Aggression, fear and sociability in everyday life of Swedish dogs. *Proc World Congr Genet Appl Livest Prod Species Breeding: Companion Animals (Posters)* 795:1–3.
- Blackwell EJ, Twells C, Seawright A, Casey RA (2008) The relationship between training methods and the occurrence of behavior problems, as reported by owners, in a population of domestic dogs. *J Vet Behav Clin Appl Res* 3:207–217
- Boyko AR (2011) The domestic dog: man's best friend in the genomic era. *Genome Biol* 12:216
- den Braber A, Zilhão NR, Fedko IO, Hottenga J-J, Pool R, Smit DJA et al. (2016) Obsessive-compulsive symptoms in a large population-based twin-family sample are predicted by clinically based polygenic scores and by genome-wide SNPs. *Transl Psychiatry* 6:e731
- Casey RA, Loftus B, Bolster C, Richards GJ, Blackwell EJ (2014) Human directed aggression in domestic dogs (*Canis familiaris*): occurrence in different contexts and risk factors. *Appl Anim Behav Sci* 152:52–63
- Cebamos L, Gray A, Stewart I, Tenesa A (2014) Regional heritability advanced complex trait analysis for GPU and traditional parallel architectures. *Bioinformatics* 30:1177–1179
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7
- Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV et al. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinforma* 14:128
- Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun* 9:1–16.
- Deldalle S, Gaunet F (2014) Effects of 2 training methods on stress-related behaviors of the dog (*Canis familiaris*) and on the dog-owner relationship. *J Vet Behav Clin Appl Res* 9:58–65
- Dodman NH, Karlsson EK, Moon-Fanelli A, Galdzicka M, Perloski M, Shuster L et al. (2010) A canine chromosome 7 locus confers compulsive disorder susceptibility. *Mol Psychiatry* 15:8–10

- Eken Asp H, Fikse WF, Nilsson K, Strandberg E (2015) Breed differences in everyday behaviour of dogs. *Appl Anim Behav Sci* 169:69–77
- Flint J (2003) Analysis of quantitative trait loci that influence animal behavior. *J Neurobiol* 54:46–77
- Foyer P, Bjällerhag N, Wilsson E, Jensen P (2014) Behaviour and experiences of dogs during the first year of life predict the outcome in a later temperament test. *Appl Anim Behav Sci* 155:93–100
- Friedrich J, Arvelius P, Strandberg E, Polgar Z, Wiener P, Haskell MJ (2018) The interaction between behavioural traits and demographic and management factors in German Shepherd dogs. *Appl Anim Behav Sci* 67–76.
- Fujisawa TX, Nishitani S, Iwanaga R, Matsuzaki J, Kawasaki C, Tochigi M et al. (2016) Association of Aryl Hydrocarbon Receptor-Related Gene Variants with the Severity of Autism Spectrum Disorders. *Front Psychiatry* 7:1–8.
- Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) ASReml User Guide Release 3.0. VSN International Ltd: Hemel Hempstead, HP1 1ES, UK.
- Goddard ME, Beilharz RG (1982) Genetic and environmental factors affecting the suitability of dogs as Guide Dogs for the Blind. *Theor Appl Genet* 62:97–102
- Gray A, Stewart I, Tenesa A (2012) Advanced complex trait analysis. *Bioinformatics* 28:3134–3136
- Greenwood TA, Akiskal HS, Akiskal KK, Kelsoe JR (2012) Genome-wide association study of temperament in bipolar disorder reveals significant associations to three novel loci. *Biol Psychiatry* 72:303–310
- Hall NJ, Wynne CDL (2012) The canid genome: behavioral geneticists' best friend? *Genes Brain Behav* 11:889–902
- Haverbeke A, Laporte B, Depiereux E, Giffroy J-M, Diederich C (2008) Training methods of military dog handlers and their effects on the team's performances. *Appl Anim Behav Sci* 113:110–122
- Hayward JJ, Castelhana MG, Oliveira KC, Corey E, Balkman C, Baxter TL et al. (2016) Complex disease and phenotype mapping in the domestic dog. *Nat Commun* 7:10460
- Hsu Y, Serpell JA (2003) Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *J Am Vet Med Assoc* 223:1293–1300
- Hutcheson HB, Olson LM, Bradford Y, Folstein SE, Santangelo SL, Sutcliffe JS et al. (2004) Examination of NRCAM, LRRN3, KIAA0716, and LAMB1 as autism candidate genes. *BMC Med Genet* 5:12
- Ilska J, Haskell MJ, Blott SC, Sánchez-Molano E, Polgar Z, Lofgren SE et al. (2017) Genetic characterisation of dog personality traits. *Genet: Genet* 116:192674
- Karlsson EK, Baranowska I, Wade CM, Salmon Hillbertz NHC, Zody MC, Anderson N et al. (2007) Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat Genet* 39:1321–1328
- Knoll AT, Jiang K, Levitt P (2018) Quantitative trait locus mapping and analysis of heritable variation in affiliative social behavior and co-occurring traits. *Genes Brain Behav* 17:1–17.
- Kukekova AV, Temnykh SV, Johnson JL, Trut LN, Acland GM (2012) Genetics of behavior in the silver fox. *Mamm Genome J Int Mamm Genome Soc* 23:164–177
- Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z et al. (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44:W90–W97
- Liinamo A-E, van den Berg L, Leegwater PAJ, Schilder MBH, van Arendonk JAM, van Oost BA (2007) Genetic variation in aggression-related traits in Golden Retriever dogs. *Appl Anim Behav Sci* 104:95–106
- McGreevy PD, Georgevsky D, Carrasco J, Valenzuela M, Duffy DL, Serpell JA (2013) Dog behavior co-varies with height, body-weight and skull shape. *PLoS ONE* 8:e80529
- Mehrkam LR, Wynne C (2014) Behavioral differences among breeds of domestic dogs (*Canis lupus familiaris*): Current status of the science. *Appl Anim Behav Sci* 155:12–27
- Nagamine Y, Pong-Wong R, Navarro P, Vitart V, Hayward C, Rudan I et al. (2012) Localising Loci underlying complex trait variation using regional genomic relationship mapping. *PLoS ONE* 7:e46501
- Nagelhus EA, Ottersen OP (2013) Physiological roles of aquaporin-4 in brain. *Physiol Rev* 93:1543–1562
- O'Neill DG, Coulson NR, Church DB, Brodbelt DC (2017) Demography and disorders of German shepherd dogs under primary veterinary care in the UK. *Canine Genet Epidemiol* 4:7
- Ostrander EA, Wayne RK, Freedman AH, Davis BW (2017) Demographic history, selection and functional diversity of the canine genome *Nat Rev Genet* 18:705–720
- Passani MB, Giannoni P, Bucherelli C, Baldi E, Blandina P (2007) Histamine in the brain: beyond sleep and memory. *Biochem Pharm* 73:1113–1122
- Perlis RH, Huang J, Purcell S, Fava M, Rush AJ, Sullivan PF et al. (2010) Genome-wide association study of suicide attempts in mood disorder patients. *Am J Psychiatry* 167:1499–1507
- Riggio V, Matika O, Pong-Wong R, Stear MJ, Bishop SC (2013) Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. *Heredity* 110:420–429
- Rooney N, Bradshaw J (2014) Canine welfare science: an antidote to sentiment and myth. In: Horowitz A (ed) *Domestic dog cognition and behavior*. Springer Berlin Heidelberg, pp. 241–274.
- Rooney NJ, Cowan S (2011) Training methods and owner–dog interactions: links with dog behaviour and learning ability. *Appl Anim Behav Sci* 132:169–177
- Roth LSV, Faresjö Å, Theodorsson E, Jensen P (2016) Hair cortisol varies with season and lifestyle and relates to human interactions in German shepherd dogs. *Sci Rep* 6:1–7
- Ruefenacht S, Gebhardt-Henrich S, Miyake T, Gaillard C (2002) A behaviour test on German Shepherd dogs: heritability of seven different traits. *Appl Anim Behav Sci* 79:113–132
- Saetre P, Lindberg J, Leonard JA, Olsson K, Pettersson U, Ellegren H et al. (2004) From wild wolf to domestic dog: gene expression changes in the brain. *Mol Brain Res* 126:198–206
- Saetre P, Strandberg E, Sundgren P-E, Pettersson U, Jazin E, Bergström TF (2006) The genetic contribution to canine personality. *Genes Brain Behav* 5:240–248
- Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA et al. (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* 50:912
- Saxena R, Voight BF, Lyssenko V, Burt NP, Bakker PIW, de Chen H et al. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–1336
- Schoenebeck JJ, Ostrander EA (2014) Insights into morphology and disease from the dog genome project. *Annu Rev Cell Dev Biol* 30:535–560
- Schütze S, Orozco IJ, Jentsch TJ (2016) KCNQ potassium channels modulate sensitivity of skin down-hair (D-hair) mechanoreceptors. *J Biol Chem* 291:5566–5575
- Serpell JA, Duffy DL (2016) Aspects of juvenile and adolescent environment predict aggression and fear in 12-month-old guide dogs. *Front Vet Sci* 3:1–8.
- Svartberg K (2005) A comparison of behaviour in test and in everyday life: evidence of three consistent boldness-related personality traits in dogs. *Appl Anim Behav Sci* 91:103–128
- Tang R, Noh HJ, Wang D, Sigurdsson S, Swofford R, Perloski M et al. (2014) Candidate genes and functional noncoding variants

- identified in a canine model of obsessive-compulsive disorder. *Genome Biol* 15:R25
- Tang B, Zhou Q, Dong L, Li W, Zhang X, Lan L et al. (2019) iDog: an integrated resource for domestic dogs and wild canids. *Nucleic Acids Res* 47:D793–D800
- Tiira K, Lohi H (2015) Early life experiences and exercise associate with canine anxieties. *PLoS ONE* 10:e0141907
- Trut LN (1999) Early canid domestication: the farm-fox experiment: foxes bred for tamability in a 40-year experiment exhibit remarkable transformations that suggest an interplay between behavioral genetics and development. *Am Sci* 87:160–169
- Uemoto Y, Pong-Wong R, Navarro P, Vitart V, Hayward C, Wilson JF et al. (2013). The power of regional heritability analysis for rare and common variant detection: simulations and application to eye biometrical traits. *Front Genet* 4:1–14.
- Våge J, Wade C, Biagi T, Fatjó J, Amat M, Lindblad-Toh K et al. (2010) Association of dopamine- and serotonin-related genes with canine aggression. *Genes Brain Behav* 9:372–378
- Visscher PM, Hemani G, Vinkhuyzen AAE, Chen G-B, Lee SH, Wray NR et al. (2014) Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLOS Genet* 10:e1004269
- van der Waaij EH, Wilsson E, Strandberg E (2008) Genetic analysis of results of a Swedish behavior test on German Shepherd Dogs and Labrador Retrievers. *J Anim Sci* 86:2853–2861
- Wilsson E, Sinn DL (2012) Are there differences between behavioral measurement methods? A comparison of the predictive validity of two ratings methods in a working dog program. *Appl Anim Behav Sci* 141:158–172
- Wise AL, Gyi L, Manolio TA (2013) eXclusion: toward integrating the X chromosome in genome-wide association analyses. *Am J Hum Genet* 92:643–647
- Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88:76–82
- York RA (2018) Assessing the genetic landscape of animal. *Behav Genet* 209:223–232
- Zapata I, Serpell JA, Alvarez CE (2016) Genetic mapping of canine fear and aggression. *BMC Genomics* 17:572
- Zhou X, Stephens M (2012) Genome-wide efficient mixed model analysis for association studies. *Nat Genet* 44:821–824