An across-breed validation study of 46 genetic markers in canine hip dysplasia

Lea Mikkola¹,²,³, Kaisa Kyöstilä¹,²,³, Jonas Donner⁴, Anu K. Lappalainen⁵, Marjo K. Hytönen¹,²,³, Hannes Lohi¹,²,³† and Antti Iivanainen¹*†

Abstract

Background: Canine hip dysplasia (CHD) is a common disease, with a complex genetic background. Dogs with severe CHD sometimes also suffer from osteoarthritis (OA), an inflammatory, often painful and incurable condition. Previous studies have reported breed-specific genetic loci associated with different hip dysplasia and OA phenotypes. However, the independent replication of the known associations within or across breeds has been difficult due to variable phenotype measures, inadequate sample sizes and the existence of population specific variants.

Results: We execute a validation study of 46 genetic markers in a cohort of nearly 1600 dogs from ten different breeds. We categorize the dogs into cases and controls according to the hip scoring system defined by the Fédération Cynologique Internationale (FCI). We validate 21 different loci associated on fourteen chromosomes. Twenty of these associated with CHD in specific breeds, whereas one locus is unique to the across-breed study. We show that genes involved in the neddylation pathway are enriched among the genes in the validated loci. Neddylation contributes to many cellular functions including inflammation.

Conclusions: Our study successfully replicates many loci and highlights the complex genetic architecture of CHD. Further characterisation of the associated loci could reveal CHD-relevant genes and pathways for improved understanding of the disease pathogenesis.

Keywords: Hip dysplasia, Dog, Canine, GWAS, Validations study, Neddylation

Background

Revealing the genetic background of canine hip dysplasia (CHD) has remained one of the biggest veterinary conundrums in the past few decades. Although there has been a lot of effort to uncover risk loci and causal variants, their validation and replication has proven difficult. An inadequate sample size has been an issue in many studies but studies have also been hampered by the complexity and inaccuracy of the phenotypes, and an apparent genetic heterogeneity across breeds.

* Correspondence: antti.iivanainen@helsinki.fi
† Hannes Lohi and Antti Iivanainen are Co-senior authors.
1Department of Veterinary Biosciences, University of Helsinki, P.O. BOX 66 (Agnes Sjöbergin katu 2), 00014 Helsinki, Finland
Full list of author information is available at the end of the article

© The Author(s). 2021 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
implements the hip scoring system defined by the Fédération Cynologique Internationale (FCI) [2], and uses only few specialised veterinarians to evaluate the hip scores, therefore reducing the inter-observer bias [3].

The single nucleotide polymorphic genetic markers (SNPs) evaluated in this study have been associated with CHD by us and others over the past ten years. Zhou et al. (2010) identified a total of six SNPs in their association study of Norberg angle (a measure of hip joint laxity) and OA in several breeds [4]. Friedenberg et al. (2011) found a homozygous deletion haplotype (intronic to Fibrillin 2; FBN2) associated with a severe form of CHD in Labrador Retrievers, as well as in 14 other breeds and in a cross-bred (Labrador Retriever - Greyhound) dog cohort [5]. Pfähler and Distl (2012) conducted a genome-wide association study (GWAS) of the FCI hip score in Bernese Mountain dogs and found three significantly associated SNPs representing two different loci [6]. Fels and Distl (2014) carried out a validation study with the FCI hip score in the German Shepherd [7]. They reported three significantly associated SNPs from canine chromosomes CFA24, CFA26, and CFA34 [7]. Fels et al. (2014) uncovered novel SNPs in previously identified quantitative trait loci (QTL) and found that nine SNPs in five loci associated significantly with the FCI hip score in their German Shepherd cohort [8]. Lavrijsen et al. (2014) used GWAS and exon sequencing to identify multiple alleles associated with the FCI hip score in Dutch Labrador Retrievers [9]. Sanchez-Molano et al. (2014) also studied Labrador Retrievers using a hip score defined by the British Veterinary Association and the Kennel Club; they revealed two significantly associated QTL including several SNPs [10]. Bartolome et al. (2015) reported a genetic predictive model for CHD based on seven SNPs they found using GWAS and candidate gene approaches in Labrador Retrievers [11]. Additional markers were provided by our collaborator at Genoscoper Laboratories [12]. Our own previous case-control GWAS of the FCI hip score in German Shepherds [13] revealed loci on CFA1 and CFA9 that harboured markers with either risk or protective alleles.

This study aimed to validate the previously reported associations to better understand their significance and overall genetic heterogeneity of CHD. We selected 52 SNPs based on prior research, and carried out across- and within breeds replication studies in nearly 1600 dogs from 10 breeds. We successfully replicated five markers across breeds and identified 20 markers in different breeds, which highlights the heterogeneous genetic architecture of CHD.

Results
We genotyped 52 SNPs that have previously been associated with CHD (Additional files 1 and 2), and carried out various case-control association analyses of CHD using an independent cohort of 1607 dogs consisting of 10 breeds. After quality control 46 SNPs and 1570 dogs (751 cases and 819 controls, 666 males and 904 females) remained for our analyses. The variant data of the 46 SNPs in our cohort are available through the European Variation Archive (https://www.ebi.ac.uk/eva). These SNPs are referred to by their ssIDs in this paper (see Additional file 1). We used the FCI hip score to categorize dogs into cases (hip score C or worse on both joints) and controls (hip score A/A). Dogs with hip score B were excluded to minimize phenotype ambiguity because although considered normal the FCI hip score B represents a borderline normal phenotype. We used the Cochran-Mantel-Haenszel 2x2xK (CMH) test for the breed-stratified data in our analyses.

Four SNPs associated with CHD across breeds
The raw P-values, the empirical P-values from the permutation procedure, and odds ratios (OR) from the across-breed CMH test as implemented in PLINK [14] are shown in Table 1. Only the markers with a significant association to CHD, and which also passed a test of homogeneity of OR across the breeds (see Methods), are reported. A total of four SNPs associated significantly with CHD in the across-breed analysis (Table 1). The markers are located on CFA1, CFA14, CFA26, and CFA37. Of these, the SNP on CFA1 is from our previous study of the FCI hip score on German Shepherds [13]. The variant ss7212922135 on CFA14 was originally associated with the FCI hip score in Bernese Mountain dogs [6]. On CFA26, ss7212922151 originally associated with the FCI hip score in German Shepherds [7], as did ss7212922122 on CFA33 [8]. The variant ss7212922139 on CFA37 originally associated with OA in a multi-breed analysis [4]. In contrast to the other four markers, ss7212922139 in CFA37 was not significantly associated with CHD in any of the breed-specific analyses.

Within-breed analyses reveal additional markers
The definition of cases and controls was the same as it was in the across-breed analysis. We used two methods, basic association analysis by X2-test of allele frequencies and logistic regression, to assess breed-specific association of the SNPs to CHD. The analysis method that demonstrated a better model fit, as evaluated by visual examination of quantile-quantile (Q-Q) plots (see Additional file 3), was chosen for each breed. Logistic regression showed better model fit for five breeds and the basic association test for the other five (see Additional file 3). Some inflation of the expected versus observed P-values was observed in three breeds: Finnish Lapphund, Golden Retriever, and Labrador Retriever (see Additional file 3). All these breeds have at least partially separated breeding lines of herding/working dogs (meaning
restricted mixing of the breeding dogs between lines), which may be the source for the inflation. The within-breed analyses showed multiple significant associations for different SNPs per breed (Table 2). The number of significantly associated SNPs varied between breeds, ranging from none in Bernese Mountain dogs to six significant associations in the Labrador Retrievers (Table 2). In total, 22 SNPs on CFA1, CFA3, CFA8, CFA11, CFA12, CFA14, CFA17, CFA21, CFA24, CFA25, CFA26, CFA33, and CFA34 demonstrated significant associations to CHD (Table 2). Six SNPs were significant in more than one breed (ss212922122, ss212922215, ss212922154, ss212922155, ss212922161, ss212922163 Table 2). These six SNPs are located on CFA33, CFA26, CFA34, CFA11, CFA1, and CFA24. The marker ss212922155 on CFA11 was significant on three breeds (Table 2), ss212922126 and ss212922153 on CFA1, as well as ss212922156 and ss212922152 on CFA8 were in high linkage disequilibrium (r^2 > 0.80) and therefore these

| Breed (N in analysis) | SNPs with significant associations with CHD | Location in CanFam3.1 | Reference Allele (minor / major) | OR (95% CI) | P-values: raw / permuted |
|-----------------------|---------------------------------------------|------------------------|----------------------------------|------------|--------------------------|
| Finnish Lapphund(303) | ss212922120 CFA1: 45382633                  | A/C                    | 7.4 x 10-03                     | 1.30 [1.07–1.57] | 0.013 / 0.021 a |
| ss2129222135         | CFA14: 56572744                             | T/C                    | 0.027                           | 1.19 [1.02–1.38] | 0.013 / 0.012 a |
| ss2129222146         | CFA14: 14157064                             | T/C                    | 0.011                           | 1.26 [1.06–1.49] | 0.019 / 0.021 a |
| ss2129222151         | CFA14: 14299531                             | C/A                    | 0.030                           | 0.84 [0.72–0.98] | 0.015 / 0.019 a |

^a^X^2^-test, ^b^Logistic regression. CI confidence interval; OR odds ratio. The odds ratios and their 95% confidence intervals were calculated by PLINK using default settings in which the minor allele increases the risk when OR > 1. The reference genome used in this study is CanFam3.1. *No significant associations were observed in Bernese Mountain dogs. #Major/minor allele designation of ss212922120 is changed in Lagotto Romagnolo cohort
SNP pairs were interpreted to represent one locus each. Thus, the 22 markers associating with CHD and with OR's deviating from 1 represent 20 different loci on 13 chromosomes (Table 2).

Enrichment of the neddylation-pathway

Our across and within breed analyses highlighted altogether 21 loci on fourteen chromosomes. These loci contain hundreds of candidate genes and we wanted to understand whether they are enriched in any cellular pathways. We performed two analyses using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) searching for the dog genes/proteins [15], including all the positional candidate genes within 1 Mb of the associated SNPs. In the first analysis, no enriched pathways were detected among the 58 positional candidate genes from the across-breed study (see Additional File 4). For the second analysis, we pooled all the 272 positional candidate genes from the across and within breed studies (Additional File 4). An enrichment in the neddylation pathway (Reactome ID: CFA-8951664) was spotted by STRING with a false discovery rate of 0.0314 (12 observed genes / 220 genes in the neddylation-pathway gene set, see Fig. 1 and Table 3).

Discussion

Previous efforts by us and others have discovered tens of loci associating with CHD across breeds. However, the significance of these findings has remained vague due to the lack of proper replication because of the variability of the phenotypes, varied study approaches and inadequate sample sizes.

Our replication study in over 1600 dogs across 10 breeds with 52 reported CHD markers validates 21 loci on fourteen chromosomes. Five loci were associated across breeds. Two loci included more than one validated marker. Inclusion of more markers in the study would have made the validation of each locus more robust and possibly allowed further selection of associated loci by the presence of clusters of validated SNPs. Also,
after our analyses several interesting markers were published [16].

In total, the replicated loci include over 250 genes with an enrichment of candidate genes in a neddylation pathway. Neddylation contributes to various cellular functions, including inflammation, commonly found in CHD and OA. Collectively, these results highlight the complex genetic background of CHD and provide important insights to the significance of the common and breed-specific loci in CHD. This helps to prioritize loci for further studies to identify causal variants and suggest a novel hypothesis for the possible contribution of the affected neddylation pathway to CHD and OA.

Four loci on CFA1, CFA14, CFA26, and CFA37 associated with CHD across breeds. The variants ss7212922118 and ss7212922120 on CFA1 were reported by us to associate with CHD in German Shepherds [13]. These SNPs locate within and upstream of NADPH Oxidase 3 (NOX3), which is a catalyst for the formation of superoxides and other reactive oxygen species. NADPH oxidases are an essential part of a reaction chain that has been suggested to contribute to the initiation of articular cartilage degradation [17]. However, NOX3 is mainly expressed in the inner ear and foetal tissues, which leaves its role in CHD equivocal.

The SNP on CFA14 (ss7212922135) was originally reported to associate with CHD in Bernese Mountain dogs [6]. This SNP did not associate with CHD in our Bernese Mountain dog cohort, which may result from different case definition between the current and the earlier study. The variant ss7212922135 originated from a study cohort where all cases (N = 33) had mild CHD (FCI score C) [6], while our Bernese Mountain dog cases represented mild-to-severe CHD based on FCI hip scoring (Ncases = 33; 42 with mild (FCI score C), 40 with moderate (FCI score D), and six with severe (FCI score E) CHD). However, ss7212922135 was significant in the across-breed analysis. This SNP lies within the ninth intron of Cortactin Binding Protein 2 encoding gene (CTTNBP2) [6]. CTTNBP2 participates in brain development and the regulation of synapse organisation. Although no obvious connection was found between this gene and CHD in the original study [6], some more recent phenotype associations have been listed in the GWAS catalogue (https://www.ebi.ac.uk/gwas/home) for CTTNBP2 that are worth noting: juvenile idiopathic arthritis and idiopathic osteonecrosis of the femoral head. Furthermore, body mass index adjusted waist-hip ratio, a measure for storage fat in humans, is listed in the GWAS catalogue [18] for ST7, WNT2, ASZ1, and CFTR, all within ±1 Mb of ss7212922135. Obesity is a known environmental risk factor for hip dysplasia and OA in both dogs and humans [19], although the mechanism is still poorly understood. Finally, we want to highlight the gene encoding Wnt Family Member 2 (WNT2) ~ 401 kb downstream ss7212922135. Wnt signalling pathways have been shown to participate in joint development, cartilage maintenance homeostasis, and the development and progression of OA [20].

Variant ss7212922151 on CFA26 was the third SNP, which demonstrated significant association to CHD in the across-breed analysis. This SNP originates from an

Table 3  Enriched candidate genes in the neddylation-pathway

| Gene   | Location in CanFam3.1 | Position of the associated marker(s) within the chromosome (locus; distance from the gene) |
|--------|-----------------------|----------------------------------------------------------------------------------------|
| ASB7   | CFA3: 40365515–40415345 | CFA3: 40302288; 63.227 kb                                                                 |
| PSMC6  | CFA8: 28914302–28947854 | CFA8: 28489064; 425.238 kb                                                                |
| FBXO10 | CFA11: 53833335–53885416 | CFA11: 29008777; 23.924.558 kb                                                            |
| DCAF10 | CFA11: 54107165–54150058 | CFA11: 29008777; 24.198.388. kb                                                           |
| ASB4   | CFA14: 20733592–20794394 | CFA14: 20864104; 69.710 kb                                                               |
| ASB18  | CFA25: 47026112–47087245 | CFA25: 47983231; 895.986 kb                                                              |
| COP58  | CFA25: 47783241–47796806 | CFA25: 47983231; 186.245 kb                                                              |
| UBE2F  | CFA25: 48503866–48558950 | CFA25: 47983231; 520.635 kb                                                              |
| ASB1   | CFA25: 48834475–48870290 | CFA25: 47983231; 851.244 kb                                                              |
| FBXW8  | CFA26: 13520605–13640975 | CFA26: 14157064; 516.089 kb                                                              |
| FBXO21 | CFA26: 13722301–13769273 | CFA26: 14157064; 387.791 kb                                                              |
| WSB2   | CFA26: 14400035–14456158 | CFA26: 14157064; 282.971 kb                                                              |

*aASB1 coordinates were derived from the Ensembl database (ENSCAFG00000012461, https://www.ensembl.org). Other coordinates were retrieved from the National Center for Biotechnology Information database (NCBI)*
association study of CHD in German Shepherds and locates within the fifth intron of Kinase Suppressor Of Ras 2 encoding gene (KSR2) [7]. KSR2, a scaffolding protein in the Ras-Raf-MEK-ERK pathway, has been associated with obesity in mice and humans [21, 22]. Although the non-SMAD-dependent TGF-β/BMP signalling in the osteoblastic lineage involves MKK3/6 and p38 [23], we are not aware of any studies reporting on their interaction with KSRs. Another gene of interest in the same loci is Nitric Oxide Synthase 1 (NOS1; ~262 kb away from ss7212922151). A large variety of different phenotypes have been reported in Nos1 murine models, including abnormal skeletal and skeletal muscle phenotypes for Nos1tm1Plh-mutants [24].

ss7212922122 on CFA33 locates within the gene encoding PEST Proteolytic Signal Containing Nuclear Protein (PCNP), and was originally found to associate with CHD in German Shepherds [8]. PCNP expression is omnipresent in different tissues, and with its ubiquitination partner Np95/ICBP90-like RING finger protein (NIRF) it may be involved in a signalling pathway of cell cycle regulation and/or genome stability [25, 26]. ss7212922122 is also about 469 kb upstream from ABI3BP. ABI3BP is a collagen and glycosaminoglycan binding molecule and an extracellular matrix structural component. Moreover, an intron variant of ABI3BP (rs9828061) has been associated with joint hypermobility measurement in humans [27].

Lastly, ss7212922139 on CFA37 locates within the first intron of Neurilpin 2 (NRP2). It was the only marker that despite not being significantly associated with CHD in any particular breed was still significantly associated with CHD in the across-breed analysis. Indeed, this SNP originated from a study, which utilised association and linkage populations of multiple breeds and their crosses [4]. In our study the marker had OR < 1, indicating protective effect. The original study did not report ORs, and the SNP associated with OA [4]. Although the current study data did not contain the direct OA phenotypes, OA is nevertheless accessed and considered when determining the FCI score. Considering the original and current study, ss7212922139 could represent a genuine “across-breed locus” for CHD and OA. Zhou et al. (2010) suggested Par-3 Family Cell Polarity Regulator Beta (PARD3B) as a candidate gene due to its association with OA of the knee in humans [4, 28]. In addition to PARD3B, NRP2 could also be a plausible candidate for OA. Neurilpin 2 is as a co-receptor for vascular endothelial growth factors (VEGFs) [29]. Increased VEGF expression has been indicated to associate with increased severity of OA [30], and a functional study demonstrated that VEGF injections into knee joints of mice induced OA [31]. Moreover, VEGF and its receptors have been studied as targets for treatment of OA [32], and an experimental study of a VEGF antibody (rhu-Mab-VEGF; Bevacizumab or Avastin®) has shown that Bevacizumab could offer a potential therapy for OA [33]. However, it remains unknown how NRP2-VEGF interaction could affect the development and progression of OA in dogs.

The current within-breed analyses revealed a multitude of loci that associated with CHD, strengthening the hypothesis that CHD has a complex genetic architecture and distinct genetic backgrounds in different breeds. The highest number of associated SNPs was found with Labrador Retriever (6), while none of the tested markers associated with CHD in our Bernese Mountain dog cohort. The associated markers and loci varied between breeds. Some SNPs demonstrated association to CHD in more than one breed and ss7212922151 on CFA26 was also significant in the across-breed analysis. We want to acknowledge here that the results in the Finnish Lapphund, Golden Retriever and Labrador Retriever breeds are inflated due to possible population stratification (as evidenced by the Q-Q plots, see Additional file 3). These breeds have both show and working/herding lines, which we could not account for in this study due to the missing line information. Thus, the breed-specific results for these three breeds should be regarded with some caution.

Some of the significant markers from the within-breed analyses had notably higher or lower ORs (Tables 1 and 2), which is probably due to higher across-breed variation for these markers. Also, worth noting was that some breeds had ORs of over 5 for certain markers (Table 2), which indicates relatively strong association to the disease outcome for such a complex disorder. Future studies should concentrate on these loci for breeds such as the Great Dane that had OR of 5.07 for ss7212922133 on CFA14, and had no marked inflation observed (see Additional file 3, plots M and N).

This study replicated altogether 21 loci with over 250 potential candidate genes raising an interesting question of the possible relationship and enrichment of the candidate genes in the associated loci predisposing to CHD. The STRING analysis revealed that the candidate genes were enriched in a single pathway, neddylation, which is a ubiquitination-like conserved post-translational protein modification process [34]. The key actor in neddylation is NEDD8, which is conjugated to its substrates with the help of enzymes E1 (activation of NEDD8; NAE1-UBA3 heterodimer), E2 (conjugation; UBE2M or UBE2F), and E3 (ligase; RBX1 or RBX2) [35]. Neddylation modifies the biochemical properties of the target substrates [34], such as many members of the cullin-family, p53, or EGFR [35, 36]. Neddylation is essential for cell cycle progression [37, 38], and it has been linked to many pathologies, especially to human cancers [36, 37].
Neddylation was recently linked to inflammatory arthritis via increased NF-κB activation, although increased expression of neddylation-related genes (NEDD8 and CUL1) were observed only in the synovium of the rheumatic arthritis and not in the controls with non-inflammatory OA [39].

Interestingly, searching the STRING database with the 12 neddylation pathway associated candidate genes from the current study (Table 3) and with 14 genes (CYBA, MAPK14, MMP2, MMP9, NCF1, NCF2, NCF4, NOX3, NOX1, NTN1, RAC1, RAC1, TRIO, VCAM1) highlighted in our previous studies on CHD on German Shepherds [13, 40] produced two gene clusters that shared 12 genes associated with Class I MHC mediated antigen processing & presentation (R-CFA-983169, https://version-11-0b.string-db.org/cgi/network?networkId=bPwtieGk6WuV, and Additional file 5).

There is cumulating evidence that inflammatory mechanisms are active in OA [41, 42]. Neddylation is demonstrated to participate in regulation of i.a. T-cell and macrophage functions during inflammation [39, 43, 44]. Both T-cell and macrophage mediated inflammatory responses have been observed in OA [42, 45], and therefore, the possible role of neddylation-pathway in the immune base of OA should be explored further.

Conclusions

We replicate 21 previously reported CHD-associated loci on fourteen chromosomes and identify neddylation as a novel candidate pathway for CHD and OA. We identify common and breed-specific loci and highlight the complex genetic architecture of CHD. Identification of the causal genes and variants in the associated loci remains as an important future task to better understand the molecular pathogenesis of CHD and its traits towards improved treatment and diagnostic options.

Methods

Study cohorts and phenotype

We used SNP genotyping to validate 52 SNPs on several chromosomes in a large cohort comprising of ten breeds. Our cohort consisted of 1607 dogs with FCI hip scores (Table 4), which were used to categorize dogs into cases (hip score C or worse on both joints, $N = 772$) and controls (hip score A/A, $N = 835$). It is worth noting that the FCI scoring does not represent a quantitative phenotype. Furthermore, the distribution of phenotype categories of the cohort is very skewed and does not meet the assumptions required for the analysis of datasets in the ordinary scale. Because of these reasons, only case-control analysis was possible. To minimize phenotype ambiguity, dogs with hips scored B were excluded. Although score B is considered normal, it nevertheless represents a borderline normal phenotype.

The participating breeds were (in order of number of samples per breed): Finnish Lapphund, Golden Retriever, Lagotto Romagnolo, Bernese Mountain dog, Samoyed, Spanish Water dog, Great Dane, Labrador Retriever, Karelian Bear dog, and Finnish Hound. These breeds were chosen for the project because the prevalence of CHD in them is at least moderate. FKC collects FCI hip scoring data into an open-access breeding database from which the phenotypes were gathered for this study [2, 46]. We investigated the prevalence of CHD in the above-mentioned breeds as the mean of observed yearly prevalences including all cases from FCI hip score C to E, measured in an eleven-year period (2006–2016, year of birth) [2, 46]. The lowest mean prevalence was observed for Labrador Retriever (19%) and the Finnish Hound (28%) [46]. Finnish Lapphund, Golden Retriever, Samoyed, Spanish Water dog, and Great Dane all had a mean prevalence of CHD between 30 and 40% during this time period [46]. The highest mean prevalence was

| Breed                          | Number of dogs before breed-specific QC (N cases/N controls) | Number of dogs after breed-specific QC (N cases/N controls) | SNPs after breed-specific QC |
|-------------------------------|----------------------------------------------------------|----------------------------------------------------------|----------------------------|
| Finnish Lapphund              | 307 (155/152)                                            | 303 (153/150)                                            | 48                         |
| Golden Retriever              | 244 (84/160)                                             | 236 (79/156)                                             | 47                         |
| Lagotto Romagnolo             | 200 (100/100)                                            | 194 (96/98)                                              | 48                         |
| Bernese Mountain dog          | 172 (94/78)                                              | 164 (89/75)                                              | 48                         |
| Samoyed                       | 150 (72/78)                                              | 149 (71/78)                                              | 48                         |
| Spanish Water dog             | 114 (56/58)                                              | 114 (56/58)                                              | 49                         |
| Great Dane                    | 114 (57/57)                                              | 113 (57/56)                                              | 50                         |
| Labrador Retriever            | 102 (55/47)                                              | 100 (54/46)                                              | 48                         |
| Karelian Bear dog             | 102 (49/53)                                              | 97 (47/50)                                               | 49                         |
| Finnish Hound                 | 102 (50/52)                                              | 102 (50/52)                                              | 49                         |
| Total quantity                | 1607 (772/835)                                           | 1572 (753/819)                                           | –                          |
observed for the Karelian Bear dog (41%), Bernese Mountain dog (43%), and Lagotto Romagnolo (45%) [46].

To minimise possible genomic stratification and subsequent inflation of the test statistics we were careful not to incorporate close relatives and included only one individual from all core-families during the initial data collection. Nevertheless, more distant relatedness may exist within the breeds. Also, Finnish Lapphund, Golden Retriever and Labrador Retriever are breeds that have working/herding breeding lines, which are at least partially separate from the rest of the breeding dogs. In the current study we did not have the breeding line information for these dogs and could not account for it in the analyses. This may cause additional inflation of the test statistics due to stratification within the breed. The breed-specific results for these three breeds must therefore be interpreted with some caution.

SNP genotyping
Agena MassARRAY® iPLEX was used to genotype 52 SNPs (Additional files 1 and 2) in 1607 dogs from ten breeds. Six of the SNPs were chosen for the project from our own study in German Shepherds [13]. Ten markers from CFA1, CFA5, CFA8, CFA20, and CFA25 were selected based on investigations by Wisdom Health. These markers are included in a patent (Patent no.: US10150998B2 [12]) and five of them were also in Lavrijsen et al. (2014) [9]. The rest of the markers were chosen from earlier studies of Zhou et al. (2010) [4], Friedenberg et al. (2011) [5], Pfahler and Distl (2012) [6], Bartolome et al. (2015) [11], Fels and Distl (2014) [7], Fels et al. (2014) [8], and Sanchez-Molano et al. (2014) [10] to evaluate whether results were replicable in our data set. Data from the previous studies (breed, size of cohort and reported raw and corrected \( \beta \)-values) are summarized for each marker in Additional file 1.

The genotyping was performed by the Institute of Molecular Medicine Finland (FIMM) Technology Centre, University of Helsinki. The genotyping was done in two separate batches; the first batch included samples from breeds Labrador Retriever, Golden Retriever, and Bernese Mountain Dog; the second batch included samples from breeds Spanish Water Dog, Karelian Bear Dog, Lagotto Romagnolo, Finnish Hound, Samoyed, Finnish Lapphund, and Great Dane. Initial quality control of data was carried out in FIMM. In the first batch one sample was discarded due to success rates lower than 70%, and four SNP assays were discarded due to unreliable or no results. In the second batch, five samples were discarded due to success rates lower than 70%, and two SNP assays were discarded due to unreliable or no results. The resulting data were delivered to us as map and ped files. The map file was based on the CanFam3.1. reference, Annotation Release 104.

Quality control
We carried out the quality control in stages for within-breed and across-breed association analyses. Initially there were 1607 samples and 52 SNPs before the QC steps. The QC thresholds for the breed-specific data were: 0.90 for per ID and per SNP call rates, 0.05 for the minor allele frequency (MAF), and 0.0001 for the cut-off \( p \)-value in check for Hardy-Weinberg Equilibrium (HWE) (done in controls only). The resulting breed-specific data is described in Table 4. SNP-specific failures to meet the QC criteria are presented in Additional file 1.

The QC for the across-breed data was done in two steps in PLINK. The first QC step was done at the breed-level before merging, with the following thresholds: 0.90 for per ID and per SNP call rates, 0.0001 for the cut-off \( p \)-value in check for HWE check in controls. MAF cut-off level was set to zero in this initial step, because some SNPs that might not pass the MAF threshold within a breed, might however pass it in the across-breed cohort. Subsequently, the SNPs and individuals that passed this initial QC step were merged into the across-breed data set. The final QC with the MAF cut-off threshold at 0.05 was then executed over the whole across-breed data, which left us 46 SNPs and 1572 samples. However, one Bernese Mountain dog and one Golden Retriever had hip scores of B/D (left hip/right hip) and they were excluded from the analysis, because we could not rule out the possibility of unilateral CHD induced by an injury or other environmental factor. Therefore, we finally had 1570 dogs in our analyses, of which 751 were cases, 819 were controls, and 666 were males and 904 were females.

Association analysis
Before the analyses we tested the quality-controlled SNP data for possible batch effects in R with the glm function. This was done because the genotyping was executed in two batches. We did not observe any significant batch effects. The within-breed analyses were carried out with the \(--\text{assoc}\) (X\(^2\) test of allele frequencies) or with \(--\text{logistic}\) (logistic regression) functions in PLINK, with age at radiographing as covariate in the logistic model (\(--\text{assoc}\) cannot use covariates).

As the SNPs in the study are a strongly selected and small subset of all the SNPs in the genome, the quantification of inflation is a challenge. QQ-plots were used to assess the overall fit of the alternative models and to select between them (Additional file 3). The QQ-plot is a more representative proxy for the fit of the model than the lambda value by PLINK which is based on the ratio of single (median) value of the test variable.

Some inflation of the test statistics in both analysis methods was observed in three breeds (see Additional
We used STRING analysis (V11.0) [15] to carry out a pathway analysis of the candidate gene sets, which we acquired from our association analyses. All genes from within 1 Mb from the variant that demonstrated significant association to CHD were used as an input for the STRING Multiple proteins search. The used candidate gene sets were listed in the Additional file 4. The STRING database was queried with 272 canine genes but seven genes (CNKSR3, FOXE1, MAPK14, MMP2, MMP9, NCF1, NCF2, NOX3/MMP2/MMP9/TRIO-associated genes (CYBA, MARK14, MMP2, MMP9, NCF1, NCF2, NCF4, NOX3, NOX1, NTN1, RAC1, TRIO, VCAM1) from our previous studies on German Shepherds. Genes belonging to R-CFA-983169 are marked red. For a high resolution image and the original analysis, see: https://version-11-0b.string-db.org/cgi/network?networkId=bPwtieGk6WuV.

Acknowledgements
We acknowledge the Institute of Molecular Medicine Finland (FIMM) Technology Centre for their help in the genotyping. We thank Milla Ahola for valuable discussions about model fitting and other statistical aspects of our study. We also wish to thank Sini Karjalainen for her technical assistance during the project. Finally, we are sincerely grateful for all the dog owners who donated samples from their dogs for the study.

Authors’ contributions
Design and concept: HL, AI, LM. Materials, methods and reagents: LM, HL, MKH, AL, JD. Data analyses: LM and KK. The manuscript was drafted by LM, HL, and AI, and was commented and approved by others.

Funding
This study was funded by the Academy of Finland (252602, www.aka.fi/en, to AI), The American Kennel Club Canine Health Foundation (01828, www.akcchf.org, to AI), The Aatos and Jane Erkko Foundation (jaes.fi/en, to HL), and Biocentrum Helsinki (to HL).

Availability of data and materials
The datasets generated and analysed during this study are available in the Figshare repository: https://doi.org/10.6084/m9.figshare.11369511. The data was anonymised to protect the privacy of the dog owners.

Ethics approval and consent to participate
We followed guidelines for research ethics and good scientific practice. We have an ethical license to collect EDTA blood samples (ESAVI/343/04.10.07/2016) from the Animal Experiment Board in Finland that is under the Regional State Administrative Agency of Southern Finland. The dog owners have signed a form of consent at the time of the blood sampling.
38. Tateishi K, Omata M, Tanaka K, Chiba T. The NEDD8 system is essential for cell cycle progression and morphogenetic pathway in mice. J Cell Biol. 2001;155:571–9. https://doi.org/10.1083/jcb.200104035.

39. Liu K, Chen K, Zhang Q, Zhang L, Yan Y, Guo C, et al. TRAF6 neddylation drives inflammatory arthritis by increasing NF-kB activation. Lab Invest. 2019;99:528–38. https://doi.org/10.1038/s41374-018-0175-8.

40. Mikkola L, Holopainen S, Pessa-Morikawa T, Lappalainen AK, Hytönen MK, Lohi H, et al. Genetic dissection of canine hip dysplasia phenotypes and osteoarthritis reveals three novel loci. BMC Genomics. 2019;20:1027.

41. Goldring MB, Otero M. Inflammation in osteoarthritis. Curr Opin Rheumatol. 2011;23:471–8. https://doi.org/10.1097/BOR.0b013e328349c2b1.

42. Robinson WH, Lepus CM, Wang Q, Raghu H, Mao R, Lindstrom TM, et al. Low-grade inflammation as a key mediator of the pathogenesis of osteoarthritis. Nat Rev Rheumatol. 2016;12:580–92. https://doi.org/10.1038/nrrheum.2016.136.

43. Jin H, Liao L, Park Y, Liu Y-C. Neddylation pathway regulates T-cell function by targeting an adaptor protein Shc and a protein kinase Erk signaling. Proc Natl Acad Sci U S A. 2013;110:624. https://doi.org/10.1073/PNAS.1213819110.

44. Chang F-M, Reyna SM, Granados JC, Wei S-J, Innis-Whitehouse W, Maffi SK, et al. Inhibition of neddylation represses lipopolysaccharide-induced proinflammatory cytokine production in macrophage cells. J Biol Chem. 2012;287:35756–67. https://doi.org/10.1074/jbc.M112.397703.

45. de Lange-Brokaar BJE, Ioan-Facsinay A, van Osch GJVM, Zuurmond A-M, Schoones J, Toes REM, et al. Synovial inflammation, immune cells and their cytokines in osteoarthritis: a review. Osteoarthr Cartil. 2012;20:1484–99. https://doi.org/10.1016/j.joca.2012.08.027.

46. Finnish Kennel Club. Breeding Database. 2019. https://jalostus.kennelliitto.fi/frmEtusivu.aspx?FCI=0. Accessed 23 Aug 2019.

47. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, et al. Assessing the impact of population stratification on genetic association studies. Nat Genet. 2004;36:388–93.

48. Wiberg RAW, Gaggiotti OE, Morrissey MB, Ritchie MG. Identifying consistent allele frequency differences in studies of stratified populations. Methods Ecol Evol. 2017;8:1899–909. https://doi.org/10.1111/2041-210X.12810.

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