Visualization of Genome Diversity in German Shepherd Dogs

Sally-Anne Mortlock, Rachel Booth, Hamutal Mazrier, Mehar S. Khatkar and Peter Williamson

Faculty of Veterinary Science, The University of Sydney, NSW, Australia.

Supplementary Issue: Current Developments in Veterinary Bioinformatics

ABSTRACT: A loss of genetic diversity may lead to increased disease risks in subpopulations of dogs. The canine breed structure has contributed to relatively small effective population size in many breeds and can limit the options for selective breeding strategies to maintain diversity. With the completion of the canine genome sequencing project, and the subsequent reduction in the cost of genotyping on a genomic scale, evaluating diversity in dogs has become much more accurate and accessible. This provides a potential tool for advising dog breeders and developing breeding programs within a breed. A challenge in doing this is to present complex relationship data in a form that can be readily utilized. Here, we demonstrate the use of a pipeline, known as NetView, to visualize the network of relationships in a subpopulation of German Shepherd Dogs.

KEYWORDS: German Shepherd Dog, genome, diversity

Introduction

The genetic composition of a breed determines the physical and physiological parameters that make up individuals. Studies have shown that there is a loss of the total amount of genetic polymorphism, generally referred to as genetic diversity, in modern dog populations. Genetic diversity has been affected by population bottlenecks as dogs were domesticated from the wild population, and again through selection for breed type. In particular, purebred dogs have been selectively bred and, in some lines, closely bred with the dominant use of popular sires, resulting in a reduction of genetic diversity. This is revealed by the presence of strong selection signatures when comparing dog genotypes from different breeds.[1,2] For purebred dogs, this has been influenced by morphological criteria, as defined by breed standards, and by the dominant influence of champion show dogs. Overall, the loss of diversity is estimated to be >90% compared to ancestral dog populations,[3,4] prompting calls for the use of the remaining diversity for maintaining health.[5,6]

The canine genome sequencing project has generated a greater understanding of genomic structure and variation within and between dog breeds.[7–9] It has also enabled the generation of genomic tools to measure metrics across the genome scale, providing detailed insight into genome function,[10–15] genetic diversity,[16–18] and inherited disorders.[19–25] Diversity can be measured using the calculations of population parameters such as "effective

population size," which estimate the equivalent number of individuals contributing to the breeding population that would give rise to the observed variance in gene frequency and inbreeding rate in the population.[26] The larger the effective population size, the greater the predicted diversity. With the advent of low-cost genotyping, molecular data are much more readily available and accessible to measure diversity accurately. However, devising ways to present such high content data in a digestible and effective form, so that they are easily interpreted and applied, is a challenge. One way to do this is to use a visual representation based on network analysis, which has the capacity to provide an intuitive approach to interact with complex data.[27]

The German Shepherd Dog (GSD) is the largest breed (purebred) dog population in Australia[28] and is very popular worldwide. Originally used as a herding dog, it has found widespread value as a working dog and as a pet. Using genome-wide genotyping data from a subpopulation of GSD as an example, this study employed the NetView pipeline, a bioinformatic tool developed for livestock[27] and with broad potential utility,[29] to visualize a network that represents individual relationships in a subpopulation of GSD.

Materials and Methods

Study animals. The dogs selected for this study were all purebred GSDs from Australia. Animals were identified

for the study through the German Shepherd Dog League of New South Wales (NSW) and at National Championship Shows. A total of 82 dogs (28 males and 54 females) from 50 different kennels, aged between 1 and 10 years, were included in the analysis (Supplementary Table 1). All procedures were performed in accordance with the guidelines for the use of animals in this research (Animal Research Act, NSW, Australia) and were approved by the Animal Ethics Committee of University of Sydney under protocols 444 and 4949. Owners of all dogs provided written informed consent.

DNA extraction. Blood was collected from the dogs by a trained professional using EDTA-coated vacutainers. Genomic DNA was isolated from the leukocyte fraction using a modified spin column procedure according to the manufacturer's recommendations (DNeasy or QIAmp; Qiagen). Total DNA was eluted into nuclease-free water, and an initial measurement of DNA quality and concentration was estimated by optical density measurement using spectrophotometry (GE Life Sciences/Pharmacia Biotech).

Genotyping. Genotyping was conducted on the 82 DNA samples using the CanineHD BeadChip (Illumina) and genome core services (GeneSeek Agrigenomics and Veterinary Diagnostics). The high-density bead chip provided comprehensive genome-wide coverage using 173,650 evenly spaced single-nucleotide polymorphisms (SNPs).

Homozygosity and inbreeding coefficient calculation. Quality control filtering was performed using the PLINK version 1.07 software.[30] The dataset was filtered to exclude individuals with >10% missing genotypes and retain SNPs with a minor allele frequency (MAF) >0.05 and SNPs with a 90% genotyping rate. Multilocus heterozygosity (MLH), runs of homozygosity (ROH), and inbreeding were also computed in PLINK version 1.07. ROH in the genomes of the 82 individuals were identified using the homozyg command and the parameters homozyg-window-het 1, homozyg-snp 100, homozyg-window-snp 50, homozyg-window-missing 5, homozyg-window-threshold 0.05, homozyg-kb 1000, homozyg-density 50, and homozyg-gap 100, and ROH were calculated as a proportion of the canine genome. Inbreeding coefficients for genotyped dogs were calculated based on the observed versus (Hardy–Weinberg) the expected number of homozygotes.

Effective population size. The effective population size (N_e) was estimated using the linkage disequilibrium method implemented by the program NeEstimator version 2.01.[31–33] To avoid bias from close linkage between loci, 10,000 evenly placed SNPs from across the genomes were selected for inclusion in the N_e calculation. The genome was divided into 10,000 segments and one SNP with the highest MAF selected from each segment; the lowest frequency alleles were screened out with an MAF cutoff set at 0.05.

Population structure. Fine-scale population structure was calculated and visualized using NetView, a population analysis pipeline.[27] An unsupervised network clustering procedure, Super Paramagnetic Clustering (SPC), was used

to create a fully connected population network in which individuals were clustered based on the genetic distance as calculated in PLINK from the SNP data. Complete linkage agglomerative clustering was used to produce a relationship matrix using the SNP data in PLINK (Supplementary File 1). The distance matrix was imported into SPC and run with $K = 10$ and minimum cluster size = 2. The resulting binary edge file produced was combined with the relationship matrix file to generate a weighted relationship matrix. The weighted relationship matrix was then converted into a GML file using the format conversion tool in Network Analysis Tools.[34] The clustering of individuals within the network was then visualized in Cytoscape using the "organic visualization" style setting from the software options.[35]

Results

NetView uses a pipeline approach to visualize networks constructed on the basis of genetic relationship analysis. The stages include (i) data preparation and editing, (ii) the computation of a genetic distance matrix, (iii) the construction of a network, (iv) cluster analysis to arrange the individuals within the network, and finally (v) visualization of the network.

Genotype data preparation and editing. Large-scale molecular genetic data, including SNP data, must first be processed for QC and edited to remove anomalies. We used PLINK, a publically available genome analysis toolset,[30] to identify any low call rates and duplication of samples, and to calculate allele frequencies. All 82 individual dogs had high-quality genotyping data and were included in the data analysis. From the original 173,650 SNPs present on the array, 89,265 SNPs remained following frequency and genotype pruning. Of these, 2,997 SNPs failed the missingness test (geno >0.1), while 83,641 SNPS failed the frequency test (MAF <0.05). Therefore, subsequent analysis was performed on 89,265 SNPs for all dogs.

Homozygosity and inbreeding. Based on the mean MLH value, the degree of heterozygosity in this cohort of GSD was 0.3. When considering the average of all dogs, a total of 14.16% of the genomes were covered with ROH >1 MB, with an average of 177.4 runs per dog, while only 1.61% was covered by ROH >4 MB, and 0.10% by ROH >8 MB (Table 1; Supplementary File 2). Using data from all dogs, the effective population size (N_e) was calculated to be 43. Furthermore, when the molecular data was used to assess autozygosity, the F_{ROH} for lengths >1MB averaged 0.119. Values of F_{ROH} for other defined thresholds are listed in Table 1.

Genetic distance matrix. Stratification within the 82 dogs was determined using the allele sharing distance (ASD) calculation. ASD was calculated as follows: 1 – average proportion of shared alleles. A module to help perform these calculations, based on the identity by state, is available in the PLINK toolset. The resulting distance matrix was used as an input for network computation and was visualized using a color-coded heat map. Pairwise genetic relationships between the 82 GSDs, based on the genome-scale data, are presented in Figure 1A.

Table 1. The average proportion of the genome covered in ROH and the average number of runs of different size in German Shepherds.

LENGTH* (MB)	PROPORTION OF GENOME (%)	NO. OF RUNS	AVERAGE SNPs PER	F_{ROH}
>1	14.16	177.40	98.58	0.119 ± 0.0019 (0.060−0.178)
>2	7.38	60.13	151.16	0.058 ± 0.0013 (0.026−0.103)
>4	1.61	7.66	269.80	0.011 ± 0.0006 (0.002−0.030)
>8	0.10	0.24	535.40	0.001 ± 0.0001 (0.000−0.004)

Notes: *ROH lengths >1 MB (F_{ROH} >1 MB), >2 MB (F_{ROH} >2 MB), >4 MB (F_{ROH} >4 MB), and >8 MB (F_{ROH} >8 MB) corresponding to ROHs originating from ancestral populations 50, 25, 12, and 6 generations ago, respectively.

The diagonal represents the location of each dog compared to itself and is coded to the maximum intensity of blue to indicate identity. Three prominent groupings are evident by shades of blue and in close alignment with the diagonal of the matrix. A commonly used method for visualizing similar data, a multidimensional scaling (MDS) plot, is also included (Fig. 1B).

Network construction. NetView uses unsupervised network clustering methodology, specifically SPC, implemented in the publically available software program, Sorting Points into Neighborhoods (SPIN).[36] The SPC method is based on the physical properties of ferromagnets (the magnetic phase transitions of spin systems),[37] but has been widely applied as a robust algorithm for analyzing gene expression profiles, phylogenetic trees, protein classification, and functional molecular networks.[38,39] The GSD relationship matrix was used as an input into SPIN/SPC and run with a constant K-value of 10, and a minimum cluster size of 2.

Network visualization. To visualize the network, NetView incorporates the SPIN/SPC output into the Cytoscape software. Cytoscape is a flexible open-source software platform to visualize complex biomolecular network data.[35] The resulting network is shown in Figure 2A. There are three clear clusters of closely related dogs, with 11, 12, and 14 dogs identified in each of these clusters. These clusters represent the most closely related dogs within a line. This can be compared with the relationship matrix that details the results for each of the pairwise comparisons (Supplementary Table 1). The other dogs are shown to be more distantly related, with varying contributions to the network. The population structure is represented in terms of nodes, edges between nodes, and thickness of edges. Figure 2A represents a static and dynamic display of the network, which may be explored in detail. A static example of a zoomed view of one region of the network is shown in Figure 2B. The network may also be scaled to incorporate additional genetic data or to add annotations or other types of relevant data. For example, a weighting may be included for gene coexpression data.

Discussion

Genome diversity is directly relevant to maintain a healthy breed and reduce morbidity from inherited disorders in dogs. Geneticists use measures of inbreeding and ROH as estimates of population-based descriptions of homozygosity. However, inbreeding averages do not capture risk to individuals very well and even ROH can be poorly correlated with disease prevalence, especially if based on common SNPs.[3] This may be confusing when applied to breeding programs. Visualization can reflect accurate relationships and provide a tool for mate selection in a relatively simple manner. We have adopted the tools that are described as NetView, a pipeline approach to construct a relationship network. This study demonstrates the application of NetView to visualize the genomic diversity in a cohort of GSD.

Maintenance of breed genetic diversity in dog breeds is a key goal for promoting health and welfare of animals. Different methods for measuring diversity have been developed based on genealogical and molecular data. Average values for coefficients of inbreeding are often quoted when discussing reduced diversity, and with the availability of molecular data, direct measures of genetic heterogeneity, or its converse, ROH, may be considered. ROH are classified according to the length of the observed homozygosity, with short runs reflecting more ancient haplotypes and long runs reflecting likely close relationships between parents. Long runs of ROH are enriched for deleterious mutations. Measured at a breed level, these parameters do not necessarily correlate with diseases across the entire breed, but the probability of inherited diseases arising in subpopulations of the breed is increased. Bateson and Sargan[40] estimated that, in a breed with a recessive disease allele frequency of 10%, a 13% reduction in heterozygosity translates to ~1.2% increase in the prevalence of an inherited disease. Within a breed, there is a limit to the available genetic diversity, which may be measured as the effective population size and is related to the ancestral founder population and the selective breeding effects over the intervening period. However, for purebred dogs, these practices have resulted in some very complex pedigree structures. Commonly, the complexity and reduced diversity arises from the limited use of breeding dogs, with preference given to the use of champion sires or selected dogs with desirable conformational features. As a result pronounced sire effects, such as increased inbreeding, may be restricted to subpopulations within the breed. The simplest advice for breeders has been to avoid mating dogs that are close relatives. Adopting this approach beyond the
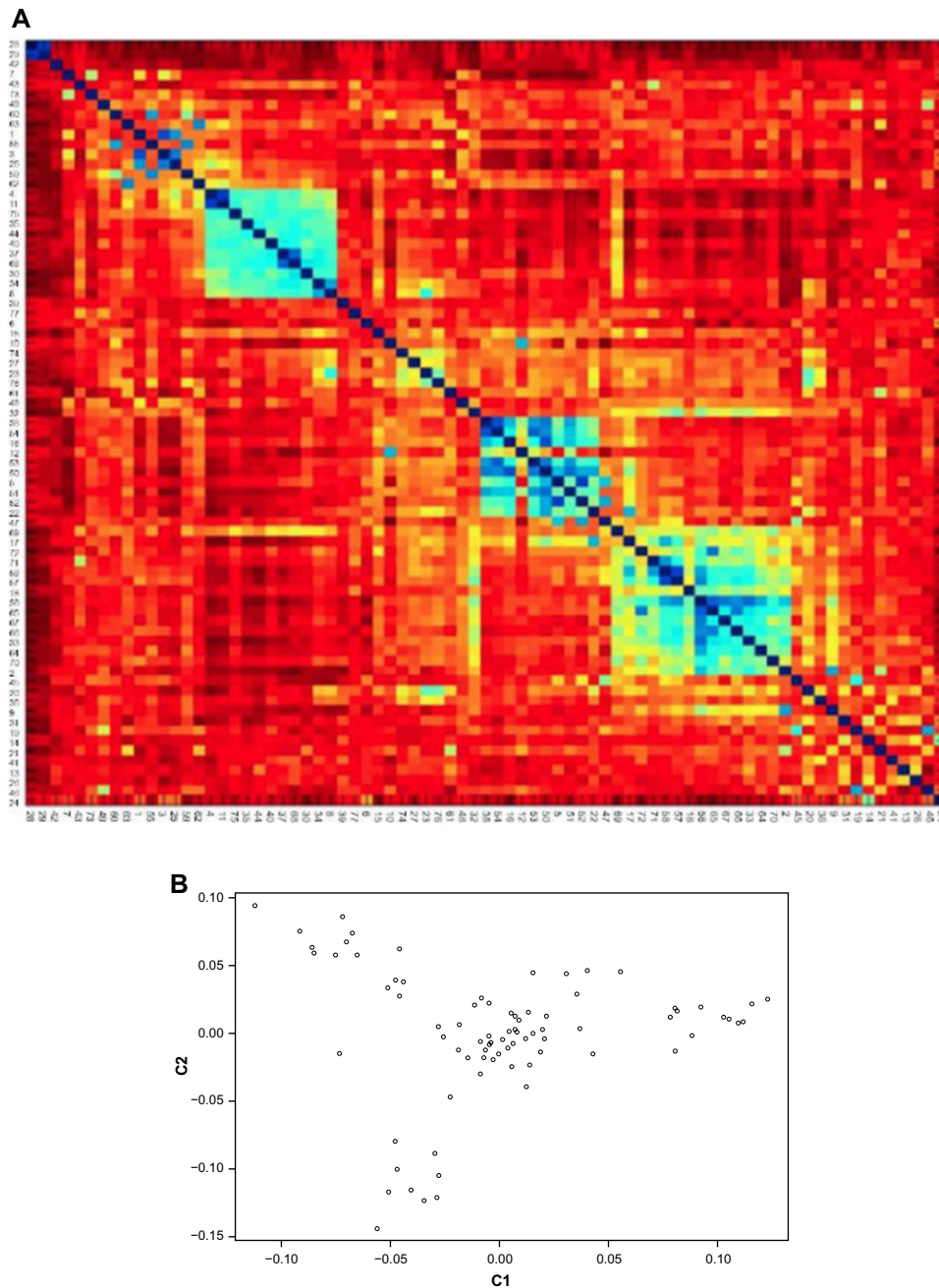
A

B

Figure 1. Representation of the genetic relationship between dogs using molecular data. (**A**) Heat map produced by NetView (Spin) depicting a relationship matrix of the GSD. Individual genetic SNP values are represented as colors contained in the matrix. The presence of three population clusters is evident in light blue coloring. (**B**) Genotyping data demonstrating relationships as represented in a two-dimensional MDS plot.

immediate family can be difficult to determine, and simply minimizing coancestry may have a little effect on the breed diversity over the longer term.[41]

With the introduction of increasing levels of available molecular data for individual dog genomes, either as a complete sequence or genome-wide SNP genotyping, the potential to identify disease-causing mutations has grown markedly. This information is now incorporated into breed health strategies for registered dogs and is helping in the decision of managing allele frequencies in the population via DNA testing and avoiding carrier matings, or even breeding to eliminate the mutation

entirely, although the latter approach needs to be carefully considered to offset potential loss of diversity.[6] With the expansion of molecular data, the capacity to accurately calculate relationships is also markedly enhanced. These data provide another level of information for breeders and hold the potential for improved choice of mating pairs when compared with simply not mating close relatives. Molecular data provide the means to calculate the actual genetic similarity of dogs, independent of geographic location and the extent of pedigree information. However, simplifying the data to present an overview of complex relationship or pedigree information is a challenge. One
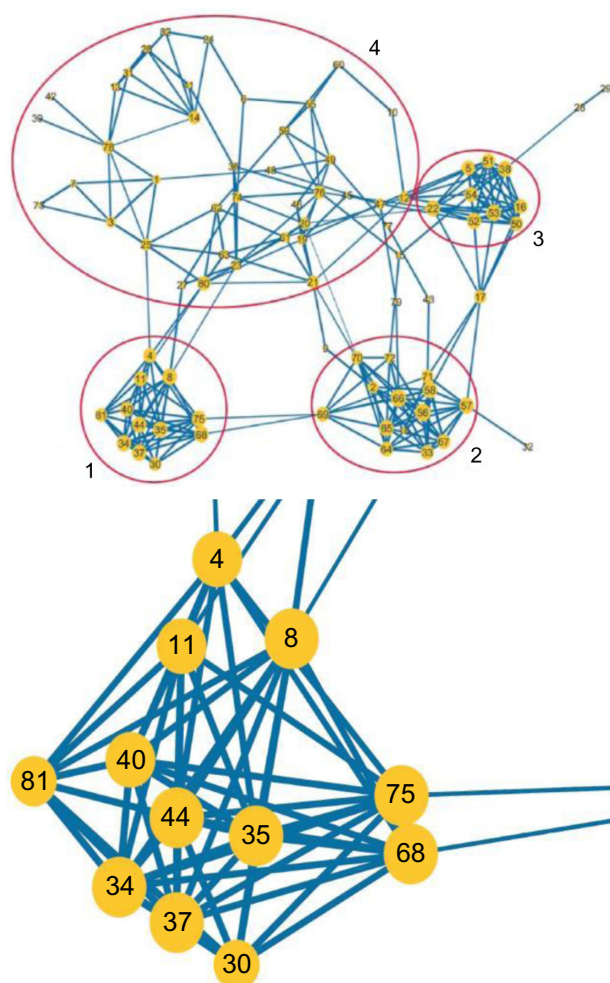
Figure 2. Network visualization graph produced by Cytoscape depicting the genetic relationships and degree of relatedness between individual GSDs based on SNP data. Individuals are represented as nodes and lines between dogs represent relationships. (**A**) Three population clusters are evident due to colocation and thicker lines between related individuals, proportional to genetic distance. (**B**) An expanded view of one cluster showing the detail of individual relationships.

approach has been adopted to develop computer-based algorithms. Windig and Oldenbroek[41] developed a computer-based simulation program to manage inbreeding at a national level for Golden Retrievers in the Netherlands. Others have developed breeding values for complex traits using traditional methods of estimated breeding values to incorporate into breeding decisions for specific disorders.[42–45] Such programs need to have oversight, since selection based on individual traits can lead to unintended adverse consequences. Despite this potential drawback, the availability of a single numerical measure of a complex trait has proven to be highly valuable for selective breeding in livestock species. However, when choosing mates simply to maintain diversity, a simple visual method to identify the degree of genetic similarity may also be beneficial, such as a network view, as demonstrated in this study.

Visualization reflects an accurate measure of the relationship between two animals from molecular data. Typically,

this creates a complex matrix or analysis that is difficult to represent in a simple format. Visualization of complex data has the advantage of providing an immediate source of information on the diversity spectrum across a subpopulation or even across the entire population of extant dogs within a breed. Interpretation of the visualized data is intuitive, and the network may be presented in a dynamic form to allow for exploration of an additional level of detail. The NetView approach has the advantage of providing a pipeline protocol using free software modules. The application of the pipeline to the GSD data in this article demonstrates the step-by-step procedures. Now, this analysis pipeline has also been implemented in Python.[29]

Acknowledgments

Author Contributions

Conceived and designed the experiments: SM, HM, MK, PW. Analyzed the data: SM, RB, MK, PW. Wrote the first draft of the manuscript: SM, RB, PW. Contributed to the writing of the manuscript: SM, RB, HM, MK, PW. Agree with the results and conclusions: SM, RB, HM, MK, PW. Jointly developed the structure and arguments for the paper: SM, RB, MK, PW. Made critical revisions and approved final version: SM, RB, MK, PW. All authors reviewed and approved of the final manuscript.

Supplementary Materials

Supplementary File 1. Relationship matrix for all dogs used in this study.

Supplementary File 2. Summary of homozygous regions showing ROH ≥1 MB.

Supplementary Table 1. Sex and age of genotyped dogs.

REFERENCES

1. Akey JM, Ruhe AL, Akey DT, et al. Tracking footprints of artificial selection in the dog genome. *Proc Natl Acad Sci U S A*. 2010;107(3):1160–5.
2. Vaysse A, Ratnakumar A, Derrien T, et al. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet*. 2011;7(10):e1002316.
3. Jansson M, Laikre L. Recent breeding history of dog breeds in Sweden: modest rates of inbreeding, extensive loss of genetic diversity and lack of correlation between inbreeding and health. *J Anim Breed Genet*. 2014;131(2):153–62.
4. Pedersen N, Liu H, Theilen G, Sacks B. The effects of dog breed development on genetic diversity and the relative influences of performance and conformation breeding. *J Anim Breed Genet*. 2013;130(3):236–48.
5. Bateson P, Nicholas FW, Sargan DR, Wade CM. Analysis of the canine genome and canine health: bridging a gap. *Vet J*. 2013;196(1):1–3.
6. Wade CM. Inbreeding and genetic diversity in dogs: results from DNA analysis. *Vet J*. 2011;189(2):183–8.
7. Kirkness EF, Bafna V, Halpern AL, et al. The dog genome: survey sequencing and comparative analysis. *Science*. 2003;301(5641):1898–903.
8. Lindblad-Toh K, Wade CM, Mikkelsen TS, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*. 2005; 438(7069):803–19.

9. Hoeppner MP, Lundquist A, Pirun M, et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One*. 2014;9(3):e91172.

10. Buishand FO, Kirpensteijn J, Jaarsma AA, Speel EJ, Kik M, Mol JA. Gene expression profiling of primary canine insulinomas and their metastases. *Vet J*. 2013;197(2):192–7.

11. Hulanicka M, Garncarz M, Parzeniecka-Jaworska M, Jank M. The transcriptomic profile of peripheral blood nuclear cells in dogs with heart failure. *BMC Genomics*. 2014;15:509.

12. Isono T, Matsumoto T, Wada A, Suzaki M, Chano T. A global transcriptome analysis of a dog model of congestive heart failure with the human genome as a reference. *J Card Fail*. 2012;18(11):872–8.

13. Richards KL, Motsinger-Reif AA, Chen HW, et al. Gene profiling of canine B-cell lymphoma reveals germinal center and postgerminal center subtypes with different survival times, modeling human DLBCL. *Cancer Res*. 2013;73(16):5029–39.

14. Roy M, Kim N, Kim K, et al. Analysis of the canine brain transcriptome with an emphasis on the hypothalamus and cerebral cortex. *Mamm Genome*. 2013;24(11–12):484–99.

15. Mortlock SA, Wei J, Williamson P. T-cell activation and early gene response in Bullmastiff dogs. *PLoS One*. 2015;10:e0121169.

16. Quignon P, Herbin L, Cadieu E, et al. Canine population structure: assessment and impact of intra-breed stratification on SNP-based association studies. *PLoS One*. 2007;2(12):e1324.

17. vonHoldt BM, Pollinger JP, Lohmueller KE, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature*. 2010;464(7290):898–U109.

18. Mortlock SA, Khatkar MS, Williamson P. Comparative analysis of genome diversity in Bullmastiff dogs. dogs. PLoS One. 2016; doi 10.1371/journal.pone.0147941.

19. Schoenebeck JJ, Ostrander EA. Insights into morphology and disease from the dog genome project. *Annu Rev Cell Dev Biol*. 2014;30:535–60.

20. Meurs KM, Mauceli E, Lahmers S, Acland GM, White SN, Lindblad-Toh K. Genome-wide association identifies a deletion in the 3' untranslated region of striatin in a canine model of arrhythmogenic right ventricular cardiomyopathy. *Hum Genet*. 2010;128(3):315–24.

21. Olsson M, Frankowiack M, Tengvall K, et al. The dog as a genetic model for immunoglobulin A (IgA) deficiency: identification of several breeds with low serum IgA concentrations. *Vet Immunol Immunopathol*. 2014;160(3–4):255–9.

22. Tonomura N, Elvers I, Thomas R, et al. Genome-wide association study identifies shared risk loci common to two malignancies in golden retrievers. *PLoS Genet*. 2015;11(2):e1004922.

23. Fletcher JL, Kondagari GS, Wright AL, Thomson PC, Williamson P, Taylor RM. Myelin genes are downregulated in canine fucosidosis. *Biochim Biophys Acta*. 2011;1812(11):1418–26.

24. Willet CE, Makara M, Reppas G, et al. Canine disorder mirrors human disease: exonic deletion in HES7 causes autosomal recessive spondylocostal dysostosis in miniature Schnauzer dogs. *PLoS One*. 2015;10(2):e0117055.

25. Wolf ZT, Brand HA, Shaffer JR, et al. Genome-wide association studies in dogs and humans identify ADAMTS20 as a risk variant for cleft lip and palate. *PLoS Genet*. 2015;11(3):e1005059.

26. Wright S. Size of population and breeding structure in relation to evolution. *Science*. 1938;87:430–1.

27. Neuditschko M, Khatkar MS, Raadsma HW. NETVIEW: a high-definition network-visualization approach to detect fine-scale population structures from genome-wide patterns of variation. *PLoS One*. 2012;7(10):e48375.

28. Shariflou MR, James JW, Nicholas FW, Wade CM. A genealogical survey of Australian registered dog breeds. *Vet J*. 2011;189(2):203–10.

29. Steinig EJ, Neuditschko M, Khatkar MS, Raadsma HW, Zenger KR. NetView P: a network visualization tool to unravel complex population structure using genome-wide SNPs. *Mol Ecol Resour*. 2016;16(1):216–27.

30. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75.

31. Waples RS, Do C. LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour*. 2008;8(4):753–6.

32. Hill WG. Estimation of effective population-size from data on linkage disequilibrium. *Genet Res*. 1981;38(3):209–16.

33. Do C, Waples RS, Peel D, et al. NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (N-e) from genetic data. *Mol Ecol Resour*. 2014;14(1):209–14.

34. Brohee S, Faust K, Lima-Mendez G, Vanderstocken G, van Helden J. Network analysis tools: from biological networks to clusters and pathways. *Nat Protoc*. 2008;3(10):1616–29.

35. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.

36. Tsafrir D, Tsafrir I, Ein-Dor L, Zuk O, Notterman DA, Domany E. Sorting points into neighborhoods (SPIN): data analysis and visualization by ordering distance matrices. *Bioinformatics*. 2005;21:2301–8.

37. Blatt M, Wiseman S, Domany E. Superparamagnetic clustering of data. *Phys Rev Lett*. 1996;76:3251–4.

38. Getz G, Levine E, Domany E. Coupled two-way clustering analysis of gene microarray data. *Proc Natl Acad Sci U S A*. 2000;97:12079–84.

39. Tornow S, HW M. Functional modules by relating protein interaction networks and gene expression. *Nucleic Acids Res*. 2003;31:6283–9.

40. Bateson P, Sargan DR. Analysis of the canine genome and canine health: a commentary. *Vet J*. 2012;194(3):265–9.

41. Windig JJ, Oldenbroek K. Genetic management of Dutch golden retriever dogs with a simulation tool. *J Anim Breed Genet*. 2015;132(6):428–40.

42. Thomson PC, Wilson BJ, Wade CM, et al. The utility of estimated breeding values for inherited disorders of dogs. *Vet J*. 2010;183(3):243–4.

43. Wilson BJ, Nicholas FW, James JW, et al. Heritability and phenotypic variation of canine hip dysplasia radiographic traits in a cohort of Australian German shepherd dogs. *PLoS One*. 2012;7(6):e39620.

44. Soo M, Sneddon NW, Lopez-Villalobos N, Worth AJ. Genetic evaluation of the total hip score of four populous breeds of dog, as recorded by the New Zealand Veterinary Association Hip Dysplasia Scheme (1991–2011). *N Z Vet J*. 2015;63(2):79–85.

45. Soo M, Worth A. Canine hip dysplasia: phenotypic scoring and the role of estimated breeding value analysis. *N Z Vet J*. 2015;63(2):69–78.