

Clarifications to the genetic differentiation of German Shepherds

Our short research report on the genetic differentiation of different breeding lines in German Shepherds has stimulated a lot of interest and discussion, and also raised some questions on various forums. Healthy criticism concerning the material, methods, analyses and conclusions of a study is a welcome part of researchers' work. Thus, we have put together some answers and clarifications to the main points of the study and the research report.

Why was the research report made?

Nearly a year ago (August 2011) the then Finnish German Shepherd Association's breeding committee posed our research group a question about the possible genetic differentiation of show and working lines in the breed: Do these lines form their own gene pools and have they as a result of breeding gradually differentiated into separate lines?

Our research group is actively pursuing several research projects relating to several different dog breeds. In these projects we investigate the genes behind different traits and diseases. In order to identify the genes, the dogs' genomes are analyzed by what is called "genome-wide analysis," where each dog's genome is tested at ~173 000 different locations. The representation of a genome received from this analysis is then compared between the cases (e.g. sick dog or dog with a particular character trait) and the controls to locate genes. As a part of this analysis we can also see the structure of the studied population (breed). The genomes of close relatives are on average closer to each other than the genomes of more distant relatives. This can be illustrated by a cluster picture, such as the one in our report (Figure 1.) Dogs whose genomes are similar to each other (similar allele frequencies) place closer to each other in the picture.

When identifying disease genes, it is very important to also know the internal structure of the studied population, since it influences how the study is organized and what methods of analysis will be used to avoid false positive results. The population structure of German Shepherds presented in the research report has been formed as part of this kind of gene research project and the committee's question.

Even though identifying disease genes is often our primary aim, it is good if our research can also answer other interesting questions concerning the breed. It is important to remember that without the volunteer support of dog owners, breeders and breed associations and their efforts to collect samples, this kind of genetic research would not even be possible.

Is the study scientifically valid?

Our research laboratory is internationally recognized and rewarded, and we only use research and analytical methods that meet international standards. Even though the genetic differentiation of German Shepherds into different lines is an important issue among breeders and other breed enthusiasts, the differentiation of populations into subpopulations as a result of strong selective breeding is not a new concept in the field of science and is not alone a very significant discovery from a scientific point of view (e.g. dogs: Chang et al 2010). Literature on the methods used in the study is readily available and we can provide you with articles if you wish to take a closer look at them.

The results presented in the research report would not be scientifically novel or interesting enough to even be offered to an international scientific journal and peer review. Today, findings related to population structure are usually published as part of a larger study. This is why this time we decided on a shorter research report for the Finnish German Shepherd Association (Saksanpaimenkoiraliitto).

Since the report was aimed at dog enthusiasts, and not for a scientific article series, it has been written as straightforwardly as possible, leaving out the more exact descriptions of methods and analyses. This does not however invalidate the scientific value of this study in any way. Since the issue of breed structure interests dog people, we are planning on writing a more extensive and thorough article on the subject, with examples from different breeds, to for example the dog magazine *Koiramme*.

What about the small sample size?

The data presented in the research report consists of dogs that have taken part in our ongoing studies involving German Shepherds. The data was not collected specifically for the purpose of this analysis. Even though the size of the research data was not particularly large, the results indicate a clear differentiation of working and show lines. In this case, a small sample size even emphasizes the finding about differentiation. In both the working and the show lines, most of the dogs come from different breeders, and thus even a small sample size is reasonably comprehensive.

We are presently expecting new results from the hip dysplasia project on German Shepherds and those results will add to the earlier research. We are also planning to analyze a larger number of German Shepherds, a little over 500 dogs, from Finland and Germany during this autumn, which would illustrate the possible differences between the countries when it comes to population structure.

If the breed associations wish for a still larger study on population structure our research group is willing to analyze more dogs, at the expense of the breed association, in order to perform more exact population-genetic and diversity analyses. Our primary focus in a follow-up study would be to find out whether “the best” dogs could be identified from good working line and show line dogs more accurately, based on an analysis of the dog’s genome. For this kind of follow-up study we however need to cooperate closely with experts from the Finnish German Shepherd Association and the Finnish German Shepherd Training Union (Saksanpaimenkoirien koulutusunioni).

How were the dogs selected for the study?

The dogs have originally been selected for ongoing research projects, according to the criteria of each project, and not specifically for the study of population structures. The dogs were classified into lines by Keijo Kodis, a then member of the Finnish German Shepherd Association’s breeding committee. The classification into lines has also been reviewed by Juha Puurunen from the breeding committee. A dog was classified as a working-line dog if there were working dogs in the dog’s pedigree for a period of five generations. Similarly, if there were known show-line dogs in a dog’s pedigree the dog was classified as a show-line dog. If both lines could be identified in the pedigree, the dog was categorized as a mixed-line dog. Defining the mixed line is more challenging, which can also be observed in the report’s picture: the mixed-line dogs place logically in the middle of the show

line and the working line. The classification of dogs into these lines with the use of pedigrees correlated perfectly with the results of the genetic study.

Has the research material been selected from a specific breeder's dogs?

Our research data on both working-line and show-line dogs has come primarily from different breeders: in the show line 14/17 dogs and in the working line 26/30 dogs. Thus, the data has not concentrated on the dogs of a specific breeder or a group of breeders, but is a relatively comprehensive sample of the population.

Why can't the names of the dogs be released to the public?

Our research group treats the dog and owner information with absolute confidence, and we do not identify the dogs or use their names in any of our research reports. This is in compliance with the research ethics of canine as well as human studies. In addition to the classification of Keijo Kodis, Juha Puurunen, the current breeding specialist of the Finnish German Shepherd Association has reviewed the list of the 60 dogs, whose owners have not denied the release of the information to the breed association. Owners are asked about this separately on the blood sample form, which is filled out when the blood sample is given.

Are the conclusions really based on the results and does the used method yield valid results?

We have used methods in this study, which are widely used all over the world in the field of population-genetic research. Each dog's genome has been analyzed using a whole-genome analysis (Illuminan HD Canine SNP Chip 173K). This method has been used in for example articles concerning the genetic distance of different breeds (e.g. Boyko et al. 2010, von Holdt et al. 2010).

Also notable is the fact that population-genetic studies rarely reach as high a number of markers as we used in our study (~173 000 markers). Many studies on the similarity or genetic diversity of animal and vegetable populations have been done with only a few or a few dozen (3-40) markers (Eckert et al. 2008). The DNA research method we used is thus extremely reliable. The actual data analyses have been done with association analysis and population statistics and population genetics packages GenABEL (Aulchenko et al. 2007) and Pegas (Paradis 2009), belonging to the R software. See the above references for further information on the calculation formulas of the used software.

What is the significance of these results in terms of breeding?

Our research shows that the genetic structure of German Shepherds includes two identifiable subpopulations. Selective breeding has been practiced in the breed, and as a result, two genetically differing lines have been formed. Subpopulations are formed relatively quickly as a result of strong selective breeding, but the effects of it can also be undone, if so desired, by mixing the two populations.

In terms of population genetics, the division of a breed into two lines also decreases the genetic variation within the lines. In a specific line there is less variation than in the breed as a whole. Joining or mixing the lines could increase the genetic variation of both lines (joint gene pool). On a general level, the decrease in genetic variation can for example increase the amount of diseases in a breed and weaken the vitality of a breed, which in turn can result in for example smaller litters. This study

has not, however, defined the level of diversity in the breed or between the lines, and it does not take a stand on whether it is sufficient in today's dogs. Generally, protecting diversity is recommended in breeding for the preservation of health and vitality in the breed.

Selective breeding has originally most likely been a conscious decision made by breed association and breeders. There has been a need for different types of German Shepherds and combinations which have produced the desired traits have been favored over others. The traits are also affected by genes, and different genes have accumulated in different lines. This provides researchers an opportunity to specify particular traits of individual dogs from both lines and within the lines and to examine more closely what kinds of genes could be found behind different traits if studied with large enough data. In the future, this kind of genetic information could be used more in dog breeding.

What is the significance of this result?

On a DNA level two subpopulations can be identified in the breed: the show line and the working line. The significance of this finding in terms of breeding has been explained above.

Does the report say anything about the working abilities of German Shepherds?

The research results do not say anything about the working abilities of German Shepherds.

How is a good working dog defined?

If we wish to dig deeper into the genetics of working abilities, we must first define those traits and find objective and standardized methods to evaluate and measure their heredity. Defining different working traits is a challenging task and requires the efforts of researchers and experienced experts on working dogs. Good working ability in a dog is a result of several personality components, which are influenced by several genes. Our research group has also experience in finding multiple gene traits (Wilbe et al. 2009).

References:

Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, et al. (2010) A Simple Genetic Architecture Underlies Morphological Variation in Dogs. *PLoS Biol* 8(8): e1000451. doi:10.1371/journal.pbio.1000451

Chang, M, Yokoyama, J, Branson, N, Dyer, D, Hitte, C, Overall, K. & Hamilton S. (2009) Intra-breed Stratification Related to Divergent Selection Regimes in Purebred Dogs May Affect the Interpretation of Genetic Association Studies. *Journal of Heredity* 2009:100(Supplement 1):S28–S36 doi:10.1093/jhered/esp012

Eckert et al. (2008) Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology* 17, 1170–1188.

Wilbe et al. (2010) Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex *Nature Genetics* 42(3): 250-255

Von Holdt et al. (2010) Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 464(8): 898-903. doi:10.1038/nature08837898

Yurii S. Aulchenko, Stephan Ripke, Aaron Isaacs, Cornelia M. Van Duijn (2007) GenABEL : an R library for genome-wide association analysis. *Bioinformatics* (2007) 23 (10): 1294-1296.

Paradis, E. (2009) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* (2010) 26 (3): 419-420.