

## Genetic Diversity Testing for Cardigan Welsh Corgi

### Overview

The Veterinary Genetics Laboratory (VGL), in collaboration with Dr. Niels C. Pedersen and staff, has developed a panel of short tandem repeat (STR) markers that will determine genetic heterogeneity and diversity across the genome and in the Dog Leukocyte Antigen (DLA) class I and II regions for specific dog populations. This test panel will be useful to dog breeders who wish to use DNA-based testing to track and increase genetic diversity as a supplement to in-depth pedigree records. DNA-based information on genetic heterogeneity and diversity, along with genetic testing results for desired phenotypes and health traits, can aid in informing breeding decisions.

Genetic diversity testing of Cardigan Welsh Corgi is currently in the preliminary results/research phase. During this phase, we will continue to test more registered dogs to build the genetic database necessary to provide an accurate assessment of genetic diversity within the breed. This report is based on genetic data of **35** registered Cardigan Welsh Corgis from the USA (n=28), Denmark (n=3), Australia (n=2), and New Zealand (n=2). Although results reported herein are preliminary, this selection of individuals should provide a reasonable picture of genetic diversity in the breed. Allele and DLA haplotype frequencies will be updated as more dogs are tested. It is anticipated that new alleles at the 33 STR loci and additional DLA class I and II haplotypes will be identified in the future for the Cardigan Welsh Corgi, but these will tend to be of much lower incidence than those detected in this initial population. We will continue to add new alleles and haplotypes if they are found in the breed, and their respective frequencies will be updated if necessary.

### [ORDER TEST KITS](#)

#### **Results reported as:**

Short tandem repeat (STR) loci: A total of 33 STR loci from carefully selected regions of the genome were used to assess genetic heterogeneity and existing genetic diversity within an individual as well as across the breed. The alleles inherited from each parent are displayed graphically to highlight heterozygosity and genetic diversity in individuals and breed wide.

DLA haplotypes: Seven STR loci linked to the DLA class I and II genes were used to identify genetic differences in a region that regulates immune responses and self/non-self-recognition. Problems with self/non-self-recognition, along with non-genetic factors in the environment, are responsible for autoimmune disease, allergies, and susceptibility to infectious agents.

Internal Relatedness: The IR value is a measure of the genetic relatedness of an individual's parents. The value takes into consideration both heterozygosity of alleles at each STR loci and their relative frequency in the population. Therefore, IR values heterozygosity over homozygosity and uncommon alleles over common alleles. IR values are unique to each dog; two individuals from different sources may have identical IR values, but a quite different genetic makeup.

## **I. Introduction to the Cardigan Welsh Corgi**

### **A. Breed History [1-3]**

The Cardigan Welsh Corgi originated in Wales and is considered one of the oldest British dog breeds, with some historic reports describing the first ancient Corgis being brought to Wales from Central Europe approximately 3,000 years ago. The breed was named after the medieval kingdom of Cardiganshire, and the word 'corgi' (originally 'kergie') means 'dwarf dog' in ancient Celtic. The Cardigan Welsh Corgi is thought to be descended from Spitz and Teckel families, and is one of two Corgi breeds from Wales, together with the Pembroke Welsh Corgi. Between 1925 and 1934, Cardigan and Pembroke Welsh Corgis were listed as one breed by the Kennel Club in the United Kingdom and were allowed to interbreed. However, in 1934 they were officially recognized as separate breeds.

Cardigan Welsh Corgis were originally bred to act as cattle dogs in Welsh farms. Initially they were used to chase off potential predators and trespassing animals before cattle herds were moved to pasture; then, as herders (driving herds to pasture by nipping on the heels of cattle) as well as drovers (driving cattle from farms to meat markets and guarding the herd at night). Over time, and concurrently with its increased popularity, the breed started to be used for other purposes such as hunting, guarding, as pets, vermin exterminators, and in athletic events.

The first pair of breeding Cardigan Welsh Corgis, Cassie and Cadno, were brought to the United States in 1931. Cassie was known for producing high-quality offspring in England (Figure 1). In 1935, the American Kennel Club (AKC) recognized the Cardigan Welsh Corgi as an official breed. Currently, the breed is ranked 68 of 201 in popularity among the AKC registries.



**Figure 1.** 'Cassie', the first female Cardigan Welsh Corgi brought to the United States in 1931.

### **B. Appearance [1-3]**

As the Cardigan Welsh Corgi was originally bred as a heeler for cattle work, they are built low to the ground in order to nip at the cattle's heels and avoid being kicked. Their height must be between 10½ and 12½ inches at the withers when standing naturally. Cardigan Welsh Corgis have long bodies that end in a long tail; according to the breed standard, the ideal length/height ratio is 1.8:1. Their tail is low set, with a fox-like brush. Ideally, dogs should weigh between 30 and 38 pounds, whereas bitches should weigh between 25 and 34 pounds. Their ears are large and prominent in relation to the size of the dog, and must be rounded at the tip and carried erect when alert. The Cardigan Welsh Corgi's neck is moderately long and muscular, especially in males. The chest is moderately broad, and the hindquarters are well muscled but less wide than the shoulders. The breed is double-coated and the hair is dense and medium length. According to the breed standard, accepted colors are all shades of red, sable and brindle; black with or without tan or brindle points; blue merle (black and gray; marbled) with or without tan or brindle points. All other colors, or a predominantly white body, are disqualifications. Other disqualifications include blue eyes (or

partially blue eyes) in any coat color other than blue merle; drop ears; and any color nose other than solid black except in blue merles.

### C. Temperament [1-3]

The breed is extremely loyal, even-tempered, and affectionate. Cardigan Welsh Corgis are highly intelligent and active, never shy nor vicious.

## II. Genetic Diversity of Cardigan Welsh Corgi

### A. Population genetics based on 33 STR loci on 25 canine chromosomes

STR markers are highly polymorphic and have great power to determine genetic differences among individuals and breeds. The routine test panel contains 33 STRs consisting of those that are recommended for universal parentage determination for domestic dogs by the International Society of Animal Genetics (ISAG) and additional markers developed by the VGL for forensic purposes [4, 5]. Each STR locus is known to contain 7 to 29 different alleles (average of 15.4 alleles/locus) in the canine breeds tested at the VGL so far. Each breed, having evolved from a small number of founders and having been exposed to artificial genetic bottlenecks, will end up with only a portion of the total available diversity. Artificial genetic bottlenecks can include phenomena such as popular sire effects, geographic isolation, catastrophes, outbreaks of disease, and ups and downs in popularity which can lead to increases and decreases in population size. The alleles identified at each of the 33 STR loci and their relative frequencies for the 35 Cardigan Welsh Corgi individuals are listed in **Table 1**.

**Table 1.** Alleles and their frequencies for 33 STR markers in Cardigan Welsh Corgi (n=35). The allele that occurs at the highest frequency at each locus is bolded.

<b>AHT121</b>	<b>AHT137</b>	<b>AHTH130</b>	<b>AHTH171-A</b>	<b>AHTH260</b>	<b>AHTk211</b>
<b>92 (0.30)</b>	131 (0.24)	119 (0.09)	219 (0.19)	<b>238 (0.39)</b>	87 (0.17)
94 (0.09)	133 (0.07)	127 (0.11)	221 (0.31)	240 (0.14)	89 (0.09)
98 (0.14)	143 (0.09)	<b>129 (0.53)</b>	<b>225 (0.33)</b>	242 (0.13)	<b>91 (0.73)</b>
100 (0.16)	147 (0.09)	131 (0.27)	231 (0.03)	246 (0.10)	97 (0.01)
102 (0.01)	149 (0.14)		233 (0.07)	248 (0.01)	
104 (0.03)	<b>151 (0.37)</b>		237 (0.07)	250 (0.20)	
106 (0.26)				254 (0.03)	
112 (0.01)					
<b>AHTk253</b>	<b>C22.279</b>	<b>FH2001</b>	<b>FH2054</b>	<b>FH2848</b>	<b>INRA21</b>
284 (0.01)	116 (0.03)	<b>132 (0.30)</b>	152 (0.26)	232 (0.01)	91 (0.24)
286 (0.19)	118 (0.23)	136 (0.10)	<b>156 (0.47)</b>	234 (0.09)	<b>95 (0.46)</b>
288 (0.07)	<b>120 (0.29)</b>	144 (0.29)	160 (0.20)	236 (0.09)	97 (0.11)
<b>290 (0.69)</b>	122 (0.17)	148 (0.21)	164 (0.01)	<b>240 (0.81)</b>	101 (0.14)
292 (0.04)	124 (0.29)	152 (0.10)	172 (0.06)		105 (0.04)
<b>INU005</b>	<b>INU030</b>	<b>INU055</b>	<b>LEI004</b>	<b>REN105L03</b>	<b>REN162C04</b>
110 (0.17)	<b>144 (0.49)</b>	<b>210 (0.71)</b>	85 (0.26)	229 (0.04)	202 (0.16)
<b>124 (0.30)</b>	148 (0.27)	216 (0.01)	<b>95 (0.66)</b>	231 (0.09)	<b>204 (0.73)</b>
126 (0.19)	150 (0.13)	218 (0.27)	109 (0.09)	<b>235 (0.31)</b>	206 (0.03)
130 (0.20)	152 (0.11)			237 (0.10)	208 (0.09)
132 (0.14)				239 (0.06)	
				241 (0.14)	

245 (0.26)					
REN169D01	REN169O18	REN247M23	REN54P11	REN64E19	VGL0760
202 (0.11)	156 (0.01)	<b>266 (0.60)</b>	226 (0.13)	139 (0.01)	13 (0.16)
210 (0.23)	<b>160 (0.46)</b>	268 (0.16)	232 (0.01)	143 (0.06)	14 (0.27)
<b>212 (0.44)</b>	162 (0.04)	272 (0.21)	234 (0.26)	145 (0.09)	15 (0.13)
216 (0.20)	164 (0.16)	278 (0.03)	236 (0.20)	147 (0.14)	18.2 (0.01)
220 (0.01)	166 (0.01)		<b>238 (0.39)</b>	149 (0.20)	19.2 (0.06)
	168 (0.29)		240 (0.01)	<b>153 (0.34)</b>	22.2 (0.07)
	170 (0.03)			155 (0.16)	<b>23.2 (0.29)</b>
					24.2 (0.01)
VGL0910	VGL1063	VGL1165	VGL1828	VGL2009	VGL2409
12 (0.07)	9 (0.07)	19 (0.04)	16 (0.09)	<b>13 (0.96)</b>	13 (0.01)
17.1 (0.10)	<b>14 (0.51)</b>	20 (0.06)	18 (0.04)	14 (0.03)	15 (0.01)
<b>18.1 (0.56)</b>	15 (0.01)	<b>21 (0.34)</b>	<b>19 (0.56)</b>	15 (0.01)	16 (0.27)
19.1 (0.14)	18 (0.03)	24 (0.31)	21 (0.10)		<b>17 (0.53)</b>
20.1 (0.06)	19 (0.01)	25 (0.03)	22 (0.17)		18 (0.16)
21.1 (0.03)	20 (0.30)	26 (0.17)	28 (0.04)		19 (0.01)
22.1 (0.04)	21 (0.06)	28 (0.04)			
VGL2918	VGL3008	VGL3235			
12 (0.39)	14 (0.04)	13 (0.13)			
<b>13 (0.44)</b>	15 (0.04)	<b>14 (0.43)</b>			
14 (0.06)	16 (0.30)	15 (0.03)			
18.3 (0.11)	<b>17 (0.50)</b>	16 (0.23)			
	18 (0.11)	17 (0.17)			
		18 (0.01)			

The most noticeable aspect of allelic diversity in the Cardigan Welsh Corgi is the relatively low number of alleles found at each locus compared to more popular breeds. This number ranged from 3 (INU055, LEI004, and VGL2009) to 8 (AHT121, VGL0760) alleles per locus in the study cohort. Contrary to other pure dog breeds in which one allele predominates at each locus, Cardigan Welsh Corgi showed more evenly distributed allelic frequencies for the 33 autosomal STR loci. Predominance of a single allele was identified in only one marker (VGL2009), with the major allele occurring at a frequency of 96%. This indicates that the genomic region harboring this locus has been under strong positive selection since the formation of the breed and might be associated with highly valued trait(s).

### B. Assessment of population diversity using standard genetic parameters

Alleles for each of the 33 STR loci listed in Table 1 and their respective frequencies are used to determine basic genetic parameters for the population (Table 2). These parameters include the number of alleles found at each locus (**Na**); the number of effective alleles (**Ne**) per locus (i.e., the number of alleles that contribute most to genetic differences/heterozygosity); the observed or actual heterozygosity (**Ho**) that was found; the heterozygosity that would be expected (**He**) if the existing population was in Hardy-Weinberg equilibrium (i.e., random breeding); and the coefficient of inbreeding (**F**) derived from the Ho and He values.

**Table 2.** Standard Genetic Assessment of 35 Cardigan Welsh Corgis based on 33 autosomal STR loci. SE = standard error.

	Na	Ne	Ho	He	F
<b>Mean</b>	5.39	3.17	0.62	0.64	0.03
<b>SE</b>	0.25	0.18	0.03	0.03	0.02

The average number of alleles (Na) identified in this cohort was low, representing approximately 35% of alleles known to exist at each of these loci in all canids tested at the VGL (5.39 out of 15.4). This number is similar to that of breeds with low genetic diversity such as the Bernese Mountain Dog (36%), American Eskimo Dog (36%) and Flat-Coated Retriever (35%). The observed (actual) heterozygosity of this cohort was 0.62, which was slightly lower than the expected heterozygosity (He) of 0.64. This resulted in an F value slightly higher than zero (F=0.03), indicating that a small proportion of individuals were more inbred than the average Cardigan Welsh Corgi. Therefore, it appears that this cohort of Cardigan Welsh Corgis was carefully selected for maximal unrelatedness. Assuming that this population is representative of the entire breed, breeders have done a reasonable job in selecting the least related parents.

### C. Standard genetic assessment values for individual STR loci

Allele frequencies can be also used to perform a standard genetic assessment of heterozygosity at each of the 33 autosomal STR loci used in this study (**Table 3**). This provides an estimate of genetic similarities in the specific regions of the genome that are associated with each STR marker. Loci with low Ho and He values contribute the least to heterozygosity (i.e., diversity) among individuals, and are most likely associated with traits that define the breed's phenotypic standards. Conversely, loci with high Ho and He values are more genetically variable and can be associated with phenotypic variation among individuals within the breed. Moreover, the F value is a coefficient of inbreeding based on Ho and He; if these two values are equal, F=0 which means that the population is in HWE. The F value will be positive when there is a deficiency of heterozygotes (i.e., fewer heterozygotes than expected), whereas negative F values correspond to an excess of heterozygotes within the population.

The number of alleles (Na) identified in individual STR loci for this cohort ranged from 3 to 8 alleles per locus (**Table 1**), while the number of effective alleles (Ne) ranged from 1.09 to 4.87 alleles per locus. Observed heterozygosity (Ho) for an individual STR locus ranged from 0.086 to 0.83, while He ranged from 0.083 to 0.79 (**Table 3**), thus indicating a large range of genetic diversity across loci. Among the 33 STR loci investigated, 14 had F values lower than zero and 19 loci had F scores greater than zero; of the latter, 9 showed high inbreeding coefficients (F>0.1) (bolded on **Table 3**). Together, these findings indicate an excess of inbred alleles in the population.

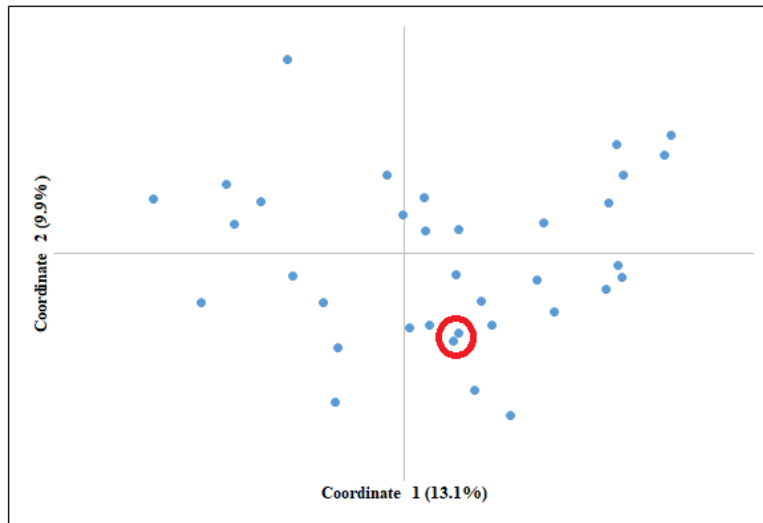
**Table 3.** Standard Genetic Assessment of individual STR loci for 35 Cardigan Welsh Corgis. Bolded loci correspond to those with high inbreeding coefficients (F>0.1).

Locus	Na	Ne	Ho	He	F
<b>AHT121</b>	8	4.767	0.829	0.79	-0.049
<b>AHT137</b>	6	4.217	0.686	0.763	<b>0.101</b>
<b>AHTH130</b>	4	2.678	0.6	0.627	0.042
<b>AHTh171-A</b>	6	3.964	0.8	0.748	-0.07
<b>AHTh260</b>	7	4.224	0.771	0.763	-0.011

<b>AHTk211</b>	4	1.761	0.486	0.432	-0.124
<b>AHTk253</b>	5	1.954	0.486	0.488	0.005
<b>C22.279</b>	5	4.07	0.8	0.754	-0.061
<b>FH2001</b>	5	4.21	0.8	0.762	-0.049
<b>FH2054</b>	5	3.014	0.743	0.668	-0.112
<b>FH2848</b>	4	1.475	0.286	0.322	<b>0.113</b>
<b>INRA21</b>	5	3.297	0.714	0.697	-0.025
<b>INU005</b>	5	4.667	0.771	0.786	0.018
<b>INU030</b>	4	2.948	0.629	0.661	0.049
<b>INU055</b>	3	1.712	0.343	0.416	<b>0.176</b>
<b>LEI004</b>	3	1.979	0.457	0.495	0.076
<b>REN105L03</b>	7	4.813	0.657	0.792	<b>0.171</b>
<b>REN162C04</b>	4	1.774	0.314	0.436	<b>0.28</b>
<b>REN169D01</b>	5	3.315	0.743	0.698	-0.064
<b>REN169O18</b>	7	3.141	0.686	0.682	-0.006
<b>REN247M23</b>	4	2.318	0.486	0.569	<b>0.146</b>
<b>REN54P11</b>	6	3.679	0.829	0.728	-0.138
<b>REN64E19</b>	7	4.685	0.771	0.787	0.019
<b>VGL0760</b>	8	4.871	0.629	0.795	<b>0.209</b>
<b>VGL0910</b>	7	2.842	0.571	0.648	<b>0.118</b>
<b>VGL1063</b>	7	2.747	0.6	0.636	0.056
<b>VGL1165</b>	7	3.945	0.771	0.747	-0.033
<b>VGL1828</b>	6	2.771	0.657	0.639	-0.028
<b>VGL2009</b>	3	1.09	0.086	0.083	-0.034
<b>VGL2409</b>	6	2.643	0.6	0.622	0.035
<b>VGL2918</b>	4	2.768	0.629	0.639	0.016
<b>VGL3008</b>	5	2.803	0.6	0.643	0.067
<b>VGL3235</b>	6	3.535	0.629	0.717	<b>0.124</b>

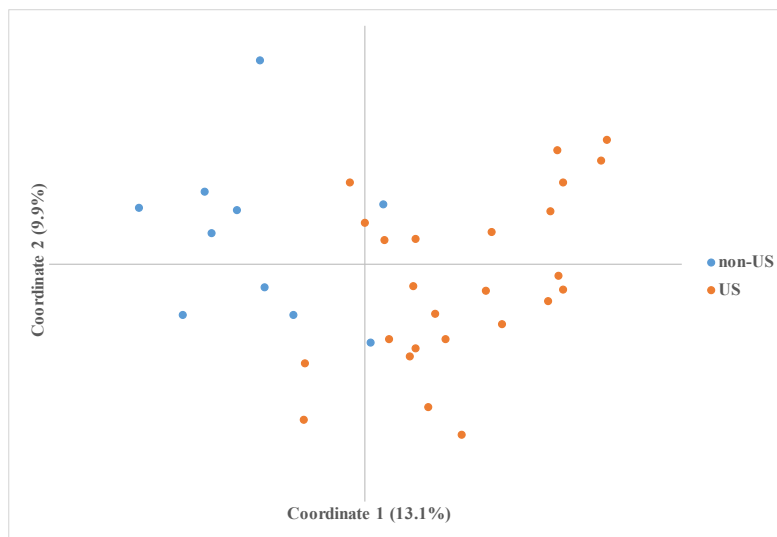
#### **D. Differences in population structure as determined by principal coordinate analysis (PCoA)**

PCoA measures the genetic relatedness of individuals in a population. The data is computed in a spherical form, but often presented in the two dimensions that most closely represent its multi-dimensional form (usually coordinates 1 and 2). The closer individuals cluster together around the XY axis, the more closely related they are to each other. The 35 Cardigan Welsh Corgis clustered as a single breed in the PCoA, with individual dogs (represented by blue dots) reasonably dispersed across all four quadrants (**Figure 2**). Some of the individuals clustered together (red circles) which suggests that they are more closely related to each other than the cohort at large. Conversely, a few dogs appeared as outliers from the main population, seen as isolated dots on the periphery of the PCoA. Overall, this cohort is comprised of individuals as unrelated as possible and is likely to represent the genetic diversity of the breed as a whole.



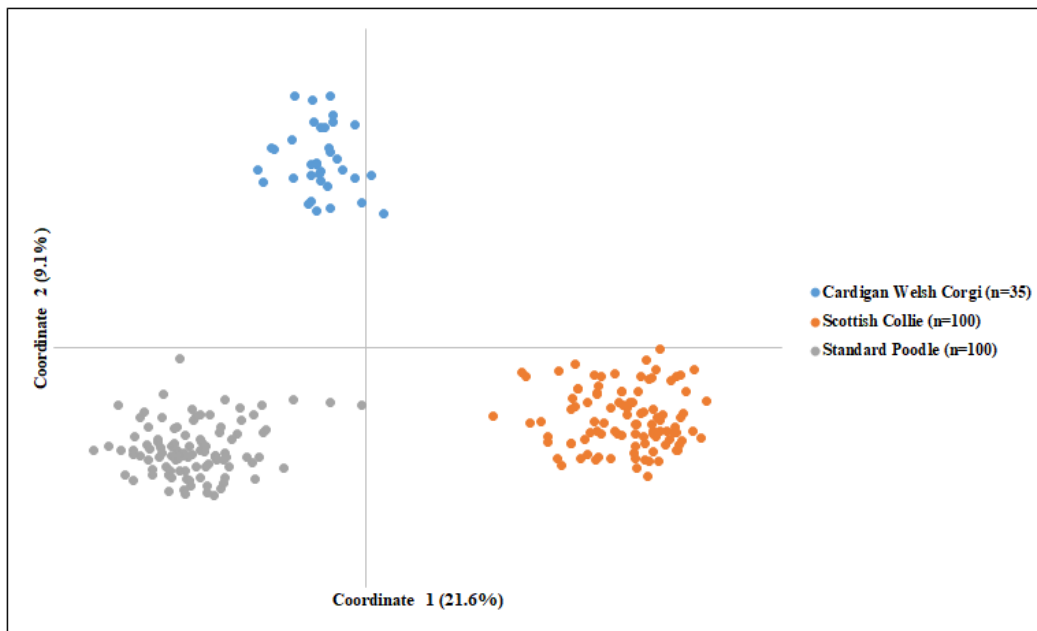
**Figure 2.** PCoA of Cardigan Welsh Corgi (n=35) based on alleles and allele frequencies at 33 autosomal STR loci. The closely related individuals are circled in red.

Additionally, the degree of relatedness among different populations of the same breed can also be visualized by using a PCoA. **Figure 3** shows the pattern of clustering of Cardigan Welsh Corgis from the US (orange dots) versus non-US dogs (blue dots) – Australia, New Zealand, and Denmark. Since the PCoA graph displays the degree of genetic differentiation between individuals, the more distant two points (representing dogs) are from each other, the greater the genetic differences between them. This analysis shows an overall *tendency* of the different populations of Cardigan Welsh Corgis (US vs. non-US) to cluster as genetically distinct varieties, represented by the orange group towards the right of the graph and the blue group towards the left quadrants. However, a few dogs cluster away from their main group and blend into the opposite variety, thus indicating a certain degree of interrelatedness among US and non-US Cardigan Welsh Corgis.



**Figure 3.** PCoA of US-born Cardigan Welsh Corgi (orange dots) and non-US dogs (blue dots) based on alleles and allele frequencies at 33 autosomal STR loci.

Finally, the degree of relatedness of individuals within the breed can be further emphasized by comparing the 35 Cardigan Welsh Corgis with a genetically distinct but somewhat closely related breed (the Scottish Collie), and a more distantly related breed (Standard Poodle) (**Figure 4**). Inter-breed clustering shows separate and well-defined populations, thus indicating that the breeds are genetically distinct as expected. However, the tighter clustering of Cardigan Welsh Corgi (blue dots) when compared to **Figure 2** can be attributed to comparisons between genetically distinct breeds, which tend to minimize minor intra-population differences.



**Figure 4.** PCoA graph comparing intra- and inter-breed relatedness of Cardigan Welsh Corgi (n=35, blue dots ) with Scottish Collie (n=100, orange dots) and Standard Poodle (n=100, grey dots).

## E. Internal relatedness (IR) of individuals and the population as a whole

### 1. IR testing

Genetic assessments such as those presented in Tables 1-3 are indicators of population-wide (mean heterozygosity), and do not reflect the genetic diversity given to individuals by their parents. Internal Relatedness (IR) is a calculation that has been used to determine the degree to which the two parents of an individual dog are related. The IR calculation takes into consideration homozygosity at each locus and gives more importance to rare and uncommon alleles. Rare and uncommon alleles would presumably be present in less related individuals. IR scores of all individuals in a population can be graphed to form a curve ranging from -1.0 to +1.0. A dog with a value of -1.0 would have parents that are totally unrelated at all 33 STR loci, while a dog with an IR value of +1.0 has parents that are genetically identical at all loci. An IR value of +0.25 would be found among offspring of full sibling parents from a random breeding population. IR values >0.25 occur when the parents of the full sibling parents are themselves highly inbred. The higher the IR value above 0.25, the more closely related are the parents and grandparents of the sibling parents.



**Table 4** summarizes the IR values for the 35 Cardigan Welsh Corgis tested. The most outbred dog in the population had an IR score of -0.21, while the most inbred dog in the group had an IR score of +0.3. This wide range of IR values shows that the degree of parental relatedness varies greatly in this cohort, a typical finding for almost all pure breeds of dogs. One-half of the Cardigan Welsh Corgis assessed in this study had IR scores equal to or greater than 0.028 and 25% of the dogs had IR scores of 0.1 or greater. This means that this cohort is comprised of a very small group of inbred dogs (IR>0.25) which is balanced by a group of strongly outbred dogs (IR scores ranging from -0.21 to -0.06). Therefore, IR scores provide a more accurate representation of heterozygosity in individual dogs than the breed-wide averages obtained from the standard genetic assessment.

**Table 4.** Internal relatedness (IR) values calculated using allele numbers and frequencies in 35 Cardigan Welsh Corgis. The IR values can be adjusted to reflect how these same dogs would score if they were to exist in a large population of village dogs (IRVD).

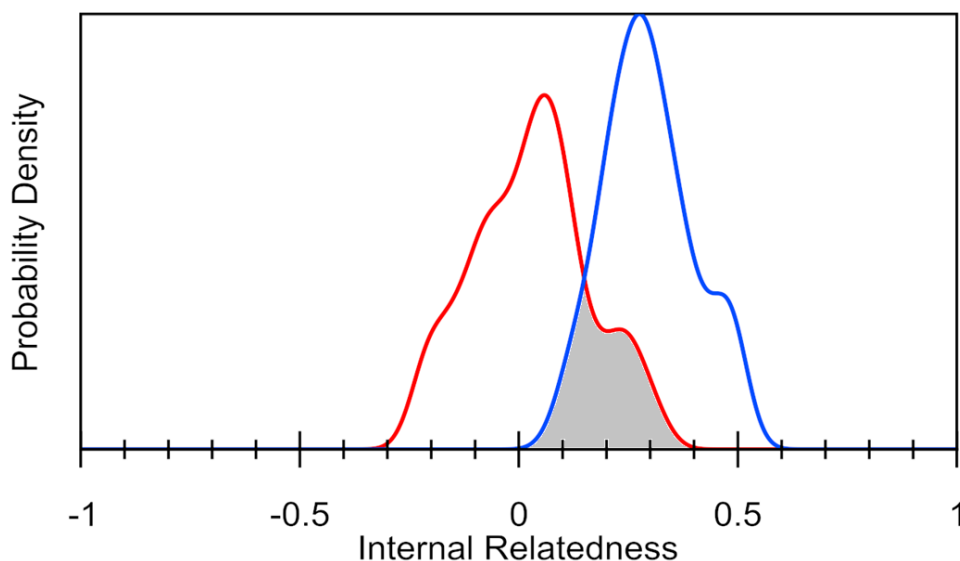
	IR	IRVD
<b>Min</b>	-0.2117	-0.0084
<b>1st Qu</b>	-0.0683	0.1839
<b>Mean</b>	0.0279	0.2304
<b>Median</b>	0.0449	0.2300
<b>3rd Qu</b>	0.1061	0.3217
<b>Max</b>	0.3047	0.4033

## 2. Adjusted IR values (IRVD) as a measure of genetic diversity lost during breed evolution.

The IR values obtained from known alleles and their frequencies can be used to approximate the amount of genetic diversity that has been lost as a breed evolves from its oldest common ancestors to the present day. Village dogs that exist throughout the SE Asia, the Middle East and the Pacific Islands are randomly breeding descendants of dogs from which most modern breeds evolved. The known alleles and their frequencies of a given breed can be compared with the same alleles and their frequency in modern village dogs to yield an adjusted IR score (IR-village dog or IRVD). Therefore, the IRVD score approximates how the IR score for a Cardigan Welsh Corgi would compare to other village dogs if its parents were also village dogs.

**Figure 4** shows that the curve representing IRVD scores for the 35 Cardigan Welsh Corgis (blue line) is shifted to the right of their actual IR scores (red line). Almost one-half of this cohort have IRVD values of 0.25 or greater (**Table 4, Figure 4**), which means that if they were found among village dogs, they would all be considered offspring of at least full sibling parents. This is a result of the relatively low amount of available village dog genetic diversity that has been retained in Cardigan Welsh Corgis.

Moreover, the gray area in **Figure 4** represents the overlap between IR and IRVD curves (25.5%), which is an estimate of the amount of genetic diversity in present-day randomly breeding village dogs that still exists in contemporary Cardigan Welsh Corgis. This amount of genetic diversity is similar to the retained genetic diversity (25%) found in all canids tested at the VGL to date (section IIB).



**Figure 4.** Distribution of IR (red line) and IR-village dog (IRVD) (blue line) values for Cardigan Welsh Corgis (n=35). The overlap between the curves (gray) represents the degree of allele sharing (25.5%) between this breed and village dogs.

#### **F. DLA class I and II haplotype frequencies and genetic diversity**

The DLA consists of four gene-rich regions that make up a small portion of canine chromosome 12. Two of these regions contain genes that help regulate normal cell- (Class I) and antibody-mediated (Class II) immunity. Polymorphisms in these regions have also been associated with abnormal immune responses, which can cause autoimmune diseases, allergies, and resistance/susceptibility to infectious diseases.

The Class I region contains several genes, but only one, DLA-88, is highly polymorphic (i.e., with many allelic forms) and is therefore most important for immune regulation. Specific alleles at the four STR loci associated with the DLA88 are linked together in various combinations, forming specific haplotypes (**Table 5**). Haplotypes are groups of genes (and consequently their alleles) inherited as a block, rather than individually.

The class II region also contains several genes, three of which are highly polymorphic: DLA-DRB1, DLA-DQB1 and DLA-DQA1. Specific alleles at these three STR loci associated with the three class II genes are strongly linked, and often inherited as a single haplotype (**Table 6**). An individual inherits one haplotype from each of the parents. The STR-based haplotype nomenclature used in this breed diversity analysis is based on numerical ranking: class I haplotypes are named 1001, 1002, and so on; class II haplotypes are named 2001, 2002, etc. It is common for different dog breeds to share common and even rare haplotypes for these loci, depending on common ancestry.

##### **1. DLA class I and II haplotypes existing in the Cardigan Welsh Corgi**

Seven DLA class I and seven DLA class II haplotypes were identified in this study cohort (**Table 5**). DLA1 haplotype 1045 was the most frequent, being identified in 37% of the dogs tested; the most frequent DLA2 haplotype was 2039 (36% frequency). The remaining DLA1 and DLA2 haplotypes were identified at lower frequencies in the population. Together, the frequencies of DLA1 haplotypes 1045 and 1012 (24%), and DLA2 haplotypes 2039 and 2003 (24%) were disproportionately high, occurring in over 60% of dogs tested. This suggests that these haplotypes

are linked, and that founders possessing these haplotypes have played an important role in maintaining predominant or desirable phenotypes of the breed; therefore, these haplotypes have been highly conserved by descent over time, a typical finding for many pure dog breeds.

**Table 5.** DLA class I and II haplotypes identified in Cardigan Welsh Corgi (n=35) with their respective frequencies. The haplotype with the highest frequency for each class is bolded.

<b>DLA I Haplotype</b>	<b>STR types</b>	<b>Frequency (%)</b>
1012	388 369 289 188	24.0
1033	382 379 277 181	9.0
<b>1045</b>	<b>376 371 277 186</b>	<b>37.0</b>
1066	376 375 277 178	10.0
1104	386 373 289 186	13.0
1207	386 371 277 186	6.0
1208	386 373 277 181	1.0

<b>DLA II Haplotype</b>	<b>STR types</b>	<b>Frequency (%)</b>
2003	343 324 282	24.0
2007	351 327 280	10.0
2017	343 322 280	13.0
2035	341 323 280	9.0
<b>2039</b>	<b>345 327 276</b>	<b>36.0</b>
2040	345 327 280	7.0
2065	339 325 286	1.0

The number of DLA class I and II haplotypes (7 and 7, respectively) found in Cardigan Welsh Corgis was amongst the lowest identified in any breed to date, and comparable to other breeds with limited genetic diversity such as the Swedish Vallhund (6/4) and Shiloh Shepherd (7/6). When compared to the diversity in DLA haplotypes of more genetically diverse breeds such as the Golden Retriever (26 and 23) and Miniature Poodle (33 and 23), this limited DLA haplotype diversity found in the breed becomes more evident. This finding suggests that a founder effect (or closely related individuals/lines) played a significant role in breed creation. However, it is possible that a few additional haplotypes will be identified as more dogs are tested (especially from outside of the US), but they are likely to occur at low frequency.

The DLA class I and II regions are frequently shared among breeds, reflecting common distant ancestry and inheritance by descent. DLA haplotype analysis showed extensive DLA I and DLA II haplotype sharing with 42 other dog breeds (**Table 6**). Interestingly, the predominant DLA I haplotype 1045 (37%) and DLA II haplotype 2039 (36%) were also the most frequent in the Collie (95% and 95%, respectively) and in the Scottish Collie (71.8% and 71.3%, respectively). This corroborates they shared distant ancestry. None of the DLA haplotypes identified in the Cardigan Welsh Corgi were exclusive to the breed.

**Table 6.** Sharing of specific DLA class I and II haplotypes between Cardigan Welsh Corgi (highlighted in blue) and other dog breeds tested at the VGL (n=42).

DLA1 #	STR types	Cardigan Welsh Corgi (n=35)	American Eskimo, Standard (n=62)	American Eskimo, Miniature (n=38)	American Akita (n=156)	Japanese Akita (n=523)	Alaskan Klee Kai (n=604)	Barbet (n=68)	Border Collie (n=60)	Berger Picard (n=111)	Bernese Mountain Dog (n=120)	Black Russian Terrier (n=145)	Biewer (n=121)	Biewer Yorkshire Terrier (n=53)	Biewer Terrier (n=195)	Yorkshire Terrier (n=16)	Borzoi (n=141)	Collie (n=49)	Doberman Pinscher (n=916)	English Bulldog (n=163)	English Mastiff (n=30)	Mastiff (n=12)	Flat Coated Retriever (n=722)	
1012	388 369 289 188	0.24	0.016	0.04	--	--	--	--	0.058	--	--	--	0.19	0.151	0.228	0.28	--	--	0.0005	0.414	--	--	--	--
1033	382 379 277 181	0.09	--	--	--	--	--	0.015	--	--	--	--	--	--	--	--	0.043	--	--	--	--	--	--	--
1045	376 371 277 186	0.37	--	--	0.003	--	--	--	0.167	--	--	--	--	--	--	--	--	0.96	0.0005	--	--	--	--	--
1066	376 375 277 178	0.1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.53	0.46	0.0007
1104	386 373 289 186	0.13	--	--	--	--	--	--	0.108	--	--	--	0.012	0.019	0.01	--	--	--	--	0.003	--	--	--	--
1207	386 371 277 186	0.06	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
1208	386 373 277 181	0.01	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

DLA1 #	STR types	Cardigan Welsh Corgi (n=35)	Golden Retriever (n=785)	Giant Schnauzer (n=291)	Havanese (n=720)	Italian Greyhound (n=1147)	Irish Red and White Setter (n=71)	Lakeland Terrier (n=124)	Labrador Retriever (n=210)	Llewellyn Setter (n=99)	Magyar Agar (n=77)	Polish Lowland Sheepdog (n=32)	Toy Poodle (n=187)	Miniature Poodle (n=346)	Poodle (n=4069)	Rat Terrier (n=39)	Samoyed (n=191)	Saint Bernard (n=71)	Scottish Collie (n=108)	Shiloh Shepherd (n=12)	Shiloh Shepherd, ISSA (n=234)	Swedish Vallhund (n=356)	Whippet (n=87)	
1012	388 369 289 188	0.24	0.0013	--	0.0146	0.0083	--	0.762	--	--	--	--	0.021	0.059	0.0145	0.05	0.013	--	0.009	--	--	--	--	0.046
1033	382 379 277 181	0.09	--	--	0.0014	--	--	--	0.002	--	0.097	--	0.004	0.0018	0.0018	--	--	--	--	--	--	--	--	--
1045	376 371 277 186	0.37	--	--	--	--	--	--	0.007	0.005	--	0.024	0.004	0.0009	--	--	--	0.718	--	0.002	--	--	0.006	
1066	376 375 277 178	0.1	0.2873	--	--	--	--	--	0.002	--	0.084	--	--	--	--	--	--	--	--	--	--	--	0.195	
1104	386 373 289 186	0.13	--	--	0.0022	--	0.032	--	--	--	0.13	--	--	--	--	--	--	--	--	--	--	--	0.144	
1207	386 371 277 186	0.06	--	--	--	--	--	--	--	--	0.013	--	--	--	--	--	--	0.007	--	--	--	--	--	
1208	386 373 277 181	0.01	--	--	--	--	--	--	--	--	0.019	--	--	--	--	--	--	--	--	--	--	--	0.029	

DLA2 #	STR types	Cardigan Welsh Corgi (n=35)	American Eskimo, Standard (n=62)	American Eskimo, Miniature (n=38)	American Akita (n=156)	Japanese Akita (n=523)	Alaskan Klee Kai (n=604)	Barbet (n=68)	Border Collie (n=60)	Berger Picard (n=111)	Bernese Mountain Dog (n=120)	Black Russian Terrier (n=145)	Biewer (n=121)	Biewer Yorkshire Terrier (n=53)	Biewer Terrier (n=195)	Yorkshire Terrier (n=16)	Borzoi (n=141)	Collie (n=49)	Doberman Pinscher (n=916)	English Bulldog (n=163)	English Mastiff (n=30)	Mastiff (n=12)	Flat Coated Retriever (n=722)
2003	343 324 282	0.24	0.016	0.04	--	--	--	0.132	0.058	--	--	--	0.231	0.208	0.246	0.38	--	--	0.0005	0.598	--	--	0.1378
2007	351 327 280	0.1	--	--	0.08	--	0.0124	--	--	--	0.083	0.041	--	--	--	--	0.152	--	--	--	--	--	--
2017	343 322 280	0.13	--	--	0.01	0.0038	--	0.029	0.175	0.126	--	--	--	--	--	--	0.007	--	--	0.215	0.22	0.17	0.0007
2035	341 323 280	0.09	--	--	--	0.3098	--	0.051	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
2039	345 327 276	0.36	--	--	0.09	0.1042	--	--	0.158	--	--	--	--	--	--	--	--	0.95	0.0093	--	--	--	--
2040	345 327 280	0.07	--	--	--	--	--	--	--	--	--	--	0.07	0.085	0.026	0.03	--	--	0.0005	--	--	--	--
2065	339 325 286	0.01	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.006	--	--	--

DLA2 #	STR types	Cardigan Welsh Corgi (n=35)	Golden Retriever (n=785)	Giant Schnauzer (n=291)	Havanese (n=720)	Italian Greyhound (n=1147)	Irish Red and White Setter (n=71)	Lakeland Terrier (n=124)	Labrador Retriever (n=210)	Llewellyn Setter (n=99)	Magyar Agar (n=77)	Polish Lowland Sheepdog (n=32)	Toy Poodle (n=187)	Miniature Poodle (n=346)	Poodle (n=4069)	Rat Terrier (n=39)	Samoyed (n=191)	Saint Bernard (n=71)	Scottish Collie (n=108)	Shiloh Shepherd (n=12)	Shiloh Shepherd, ISSA (n=234)	Swedish Vallhund (n=356)	Whippet (n=87)
2003	343 324 282	0.24	0.0217	0.033	0.2132	0.0074	--	0.762	0.014	--	--	--	0.417	0.514	0.1021	0.1	0.013	--	0.019	--	0.006	--	0.103
2007	351 327 280	0.1	0.0134	0.046	0.0465	--	0.155	--	0.04	--	--	--	0.003	0.001	0.0162	--	0.005	0.007	0.005	--	--	0.263	--
2017	343 322 280	0.13	0.0414	--	0.0076	0.2123	--	0.032	--	--	0.455	0.28	0.003	0.001	0.0021	0.03	--	--	--	0.25	0.378	--	0.247
2035	341 323 280	0.09	--	--	0.0028	0.0833	0.085	--	--	0.005	--	--	--	--	0.0001	0.01	--	0.007	--	--	--	--	--
2039	345 327 276	0.36	--	--	--	0.0833	--	--	0.007	--	--	--	--	--	0.0007	--	--	--	0.713	--	0.002	--	--
2040	345 327 280	0.07	--	--	0.0021	0.0013	--	--	--	--	--	--	0.021	0.004	0.0001	0.01	--	--	--	--	--	--	--
2065	339 325 286	0.01	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.019	--	--	--	--

## 2. Heterozygosity in the DLA region

Due to their physical proximity in canine chromosome 12, the seven loci that define the DLA class I and II haplotypes are in stronger linkage disequilibrium (i.e., have a higher probability of being inherited together) when compared to other parts of the genome measured by the 33 autosomal STR markers. However, the expectation is that these loci have achieved an equilibrium with other loci in the genome over time, and thus will also be inherited randomly. This assumption can be tested through a standard genetic assessment of each locus (**Table 7**) and averaged across all loci (**Table 8**).

**Table 7.** Standard genetic assessment for Cardigan Welsh Corgi using each of the 7 STRs in the DLA class I and II regions (n=100).

Locus	Na	Ne	Ho	He	F
<b>DLA I-3CCA</b>	4	3.04	0.66	0.67	0.021
<b>DLA I-4ACA</b>	5	3.57	0.71	0.72	0.007
<b>DLA I-4BCT</b>	2	1.88	0.51	0.47	-0.1
<b>DLA1131</b>	4	2.57	0.57	0.61	0.064
<b>5ACA</b>	5	2.95	0.63	0.66	0.049
<b>5ACT</b>	5	2.76	0.69	0.64	-0.08
<b>5BCA</b>	4	2.98	0.6	0.66	0.097

**Table 8.** Summary of standard genetic assessment for Cardigan Welsh Corgi using 7 STRs in the DLA class I and II regions (n=35).

	Na	Ne	Ho	He	F
<b>Mean</b>	4.14	2.82	0.62	0.63	0.009
<b>SE</b>	0.37	0.18	0.02	0.03	0.026

This analysis indicates that the DLA region is equally heterozygous as the rest of the genome (observed heterozygosity was estimated at 0.62 for both). The number of effective alleles (Ne), or alleles that are major determinants of heterozygosity, was slightly lower in the DLA region when compared to those observed for the 33 autosomal STR loci (2.82 and 3.17, respectively). The mean F values for autosomal STRs (0.03) and DLA STRs (0.009) were also similar and essentially zero. Together, these values indicate that heterozygosity within the DLA class I and II regions is in equilibrium with the rest of the genome.

## III. What does this assessment of genetic diversity tell us about contemporary Cardigan Welsh Corgis

Cardigan Welsh Corgis, like many other less popular breeds, have a relatively low level of genetic diversity. They lack diversity in the DLA regions, as indicated by over 60% of the individuals tested possessing the 1045/2039 and 1012/2003 DLA class I/II linked haplotypes. This is evidence that two founders, or closely related founder lines, have contributed disproportionately to the breed's evolution. This is in line with the history of the breed in the US and the fact that 80% of the 35 Cardigan Welsh Corgis used in this study are from the US. However, the fact that these haplotypes are in equilibrium with the rest of the genome indicates that this imbalance occurred at the time of breed development and has become equilibrated over the time. It is likely that additional

autosomal and DLA alleles will be identified as more dogs are tested (especially from outside of the US), albeit at lower frequency. On the other hand, the estimated amount of retained genetic diversity for Cardigan Welsh Corgi based on a comparison to indigenous dogs was 25.5%, which was similar to the average retained diversity of all breeds tested by the VGL (25%).

#### IV. Health of the Cardigan Welsh Corgi

##### A. Lifespan

Cardigan Welsh Corgis have a life expectancy of 12-15 years [1].

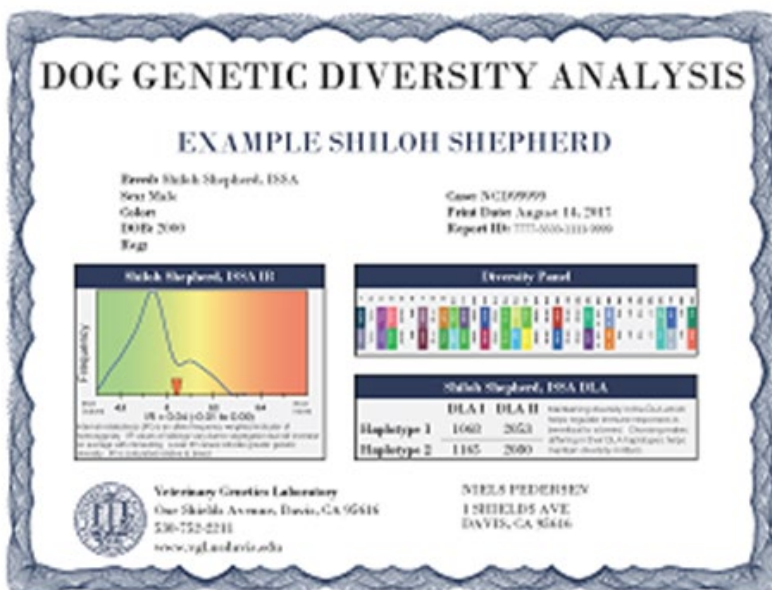
##### B. Diseases

The breed is considered very healthy in general. Breeders are advised to screen dogs for hip dysplasia, progressive retinal atrophy (PRA), and degenerative myelopathy (DM). Additionally, the Cardigan Welsh Corgi Club of America reports that canine intervertebral disc disease (IVDD) is known to occur in the breed, although the frequency has not been determined [2].

#### V. Results of VGL Canine Diversity Testing

##### A. How will you be given the results of DNA-based genetic diversity testing on your dog?

After a sample is submitted for genetic testing, the identity of the dog and owner will be replaced by a laboratory barcode identifier. This identifier will be used for all subsequent activities and each owner will be provided with a certificate that reports the internal relatedness, genomic STR genotypes and DLA class I and II haplotypes for the dog(s) tested. The internal relatedness value for the dog being tested is reported in relation to others in the population. The alleles at each of the 33 STR loci are presented as numbers that correspond to those found in Table 1. Each locus will have two alleles, which can be different (heterozygous) or the same (homozygous). Each allele is inherited from one of the parents. Dogs from closely related parents will be homozygous for more alleles at each locus, or in regions of the genome that are under strong positive selection for phenotypic trait or traits mostly favored in the breed. Dogs with a predominance of rare (i.e., low frequency) alleles will be more distantly related to the bulk of the population than dogs that have a predominance of common (i.e., high frequency) alleles. A sample genetic diversity report is shown below.



## **B. What should you do with this information?**

DNA testing for genetic diversity in the Cardigan Welsh Corgi shows a low, but acceptable, level of genome-wide genetic diversity and a relative lack of inbreeding in the study cohort. The number of DLAI and II haplotypes identified in this cohort is low compared to other dog breeds and reflects a founder effect. Therefore, the breed needs to be more carefully managed and breeders are advised to closely monitor existing diversity across the genome and in the DLA region. We believe that this can be most accurately done with DNA testing and the use of interrelatedness (IR) scores and testing for DLA I and II haplotypes to better balance and maintain genetic diversity and as a supplement to in-depth pedigrees.

The goal for breeders should be to continue to produce puppies with IR scores lower than zero, and with time, even lower scores. Mates should be preferably selected to avoid homozygosity at any genomic loci or DLA class I and II haplotype; moreover, mating of dogs with less common genomic alleles or DLA haplotypes is encouraged. Maintaining existing genomic diversity will require using IR values of potential mates based on the 33 STR loci to assure puppies of equal or greater overall diversity. However, because IR values reflect the unique genetics of individuals, they cannot be used as the primary criterion for selecting ideal mates. Mates with identical IR values may produce puppies significantly more or less diverse than their parents. Conversely, breeding dogs with high IR values (providing they are genetically different) may produce puppies with much lower IR scores than either parent. A mating between a dog with a high IR value and one with low IR value, providing the latter has few alleles and DLA haplotypes in common, will produce puppies much more diverse than the highly inbred parent. Breeders should also realize that a litter of puppies could have a wide range of IR values, depending on the comparative contributions of each of the parents. The more genetically diverse and different the parents, the greater the range of IR values in their offspring.

The next step is to compare the DLA class I and II haplotypes of the mates. You want to avoid breeding dogs that will produce puppies homozygous for the same haplotypes; once again, less common haplotypes may increase breed diversity in relation to common ones. Given the genetic diversity observed in the DLA region in Cardigan Welsh Corgi, this approach can be successful if properly implemented.

Breeders who would like to predict the genetic outcome of puppies of certain sires and dams should screen them for genetic differences in alleles and allele frequencies for the 33 genomic STR loci. Rare alleles should be favored over common ones. This information is included on all certificates and on the breed-wide data found on the VGL website.

## **VI. References**

1. American Kennel Club. Cardigan Welsh Corgi. <https://www.akc.org/dog-breeds/cardigan-welsh-corgi/>.
2. The Cardigan Welsh Corgi Club of America. <https://cardigancorgis.com/cwcca/>.
3. Wikipedia. Cardigan Welsh Corgi. [https://en.wikipedia.org/wiki/Cardigan\\_Welsh\\_Corgi](https://en.wikipedia.org/wiki/Cardigan_Welsh_Corgi).
4. Pedersen NC, Pooch AS, Liu H. A genetic assessment of the English bulldog. *Canine Genet Epidemiol.* 2016; 3:6. doi: 10.1186/s40575-016-0036-y.

5. Pedersen NC, Shope B, Liu H. An autosomal recessive mutation in *SCL24A4* causing enamel hypoplasia in Samoyed and its relationship to breed-wide genetic diversity. *Canine Genet Epidemiol.* 2017; 4:11. doi: 10.1186/s40575-017-0049-1.

**This report was generated by Felipe Avila and Shayne Hughes on 12/01/2021.**