UNRAVELING THE GENETIC CAUSE OF PATENT DUCTUS ARTERIOSUS IN THE FRIESIAN STABYHOUN

MASTER RESEARCH PROJECT

APRIL – JULY 2015

LISANNE VLUTTERS
STUDENTNR. 3516547
DATE: JULY 28, 2015

UTRECHT UNIVERSITY
FACULTY OF VETERINARY MEDICINE
SUPERVISOR: F. VAN STEENBEEK

CONTENTS

Abstract 3
Introduction 4
Aim of the study 5
Materials and methods 6
DNA samples 6
Single Nucleotide Polymorphisms 7
KASPtm Genotyping 8
Statistics 9
Results 10
Validation in Stabyhouns 10
Verification in other breeds 11
Discussion 13
Validation in Stabyhouns 13
Verification in other breeds 16
Conclusion 17
Acknowledgements 17
References 18
Appendix 1: protocol KASP™ genotyping 21

ABSTRACT

Patent ductus arteriosus is one of the most common congenital heart diseases in the dog. During fetal growth, the ductus arteriosus allows blood to flow from the pulmonary artery into the aorta so that the not-functioning lungs are bypassed. The two main factors that keep the ductus open are the low oxygen tension in the blood and the high levels of prostaglandin E_2 from the placenta. Within minutes to hours after birth, the vessel constricts and closes due to a rise in oxygen tension and the deterioration of placental PGE_2. In case of PDA, the ductus arteriosus fails to close after birth. In dogs with PDA, a very typical continuous murmur can be heard on the left heart base. Untreated, PDA will cause death within the first year in one third of all cases.

Histopathological, asymmetry and hypoplasia of the ductus muscle is seen. Also, parts of the ductus muscle contain non-contracting aorta like elastic structure. Several dog breeds are predisposed, indicating that there is a genetic component in the disease. However, in dogs, no causative gene mutations associated with the occurrence of PDA are described yet. The Friesian Stabyhoun, a Dutch breed with a small genetic base, seems to have an increased risk of PDA. Pedigree analysis revealed a high prevalence (0,55%) and inheritance (0.4). In previous association and sequence studies, 157 possible causal single nucleotide polymorphisms within two small chromosomal regions were found. The objective of this study is to make a start in elucidating gene mutations that are possibly involved in the development of PDA in the Friesian Stabyhoun.

In this research, the top 18 SNPs with the lowest p-values were screened in a group of Stabyhouns as well as a group of other breeds, using KASP™ genotyping. In the group of Stabyhoun samples (21 cases and 171 controls), a significant difference between cases and controls in allele frequency was found for two SNPs on chromosome 5. One of these SNPs is located on the PDA1 gene, which is expressed in vascular smooth muscle cells and plays a role in contractility of vascular smooth muscles. A mutation in this gene may lead to an impaired development and contractility of vascular smooth muscle cells. The second SNP is located on the PDA2 gene, a gene that is involved in maintaining the growth and absorptive function of the intestinal villi. In the group of other breeds (77 cases and 73 controls), significances were found for one SNP in allele frequency and for one SNP in genotype frequency, but they were not conclusive.

Results of this study show possible involvement of two mutations in the development of PDA in the Friesian Stabyhoun. However, PDA is a complex polygenic disease and further research is needed to find out the exact mechanism of the genetic cause of this disease. Only then, a genetic test can be developed and patients and carriers can be excluded from breeding. Additionally, the Stabyhoun dog might be a good model for other dog breeds and human, because of the small genetic base and the high prevalence of the disease in this breed.

INTRODUCTION

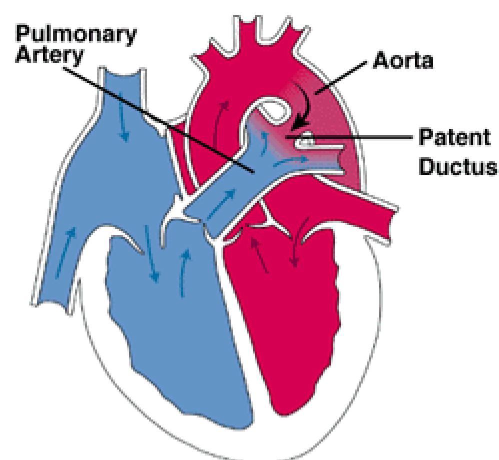

Figure 1: Patent ductus arteriosus in the heart. *During systole, oxygen rich blood from the aorta flows into the pulmonary artery which carries oxygen poor blood to the lungs.*

The ductus arteriosus, also known as ductus botalli, is a fetal vascular connection between the pulmonary artery and the aorta. It allows blood to flow from the pulmonary circulation into the general circulation, to bypass the lungs which do not function yet (Schneider 2012). During fetal development, the two main factors that keep the ductus arteriosus open are high levels of prostaglandin E_2 (PGE_2) from the placenta and the low oxygen tension in the blood (Coceani, Baragatti 2012, Dzialowski et al. 2011). After birth, opening of the lungs causes a rise in the oxygen tension in the blood, activating the endothelin signaling pathway that induces closure of the ductus arteriosus (Coceani, Baragatti 2012). Furthermore, the PGE_2- induced relaxation of the muscle deteriorates due to the loss of placental prostaglandin production (Clyman 2006, Bökenkamp et al. 2010). In healthy dogs, the ductus muscle constricts within minutes to hours after birth, closing the vessel. Within a month, total muscle degeneration is complete, leaving only the remnant elastic fibers forming the remaining ligamentum arteriosum (Buchanan 2001).

In case of patent ductus arteriosus (PDA), the ductus arteriosus fails to close after birth, leaving a 'left-to-right' shunt that allows blood to flow from the aorta into the pulmonary artery *(Figure 1).* PDA is one of the most common congenital cardiovascular abnormalities in the dog (Buchanan 1999, Kittleson, Mark D., Kienle,Richard D., 1998, Oliveira et al. 2011). The additional blood flow into the pulmonary artery causes increased pulmonary vascularity, enlargement and hypertrophy of the left atrium and left ventricle and dilatation of the aortic arch (Buchanan 2001). At clinical examination, a very typical continuous machine-like murmur can be heard on the left heart base (Buchanan, Patterson 2003). If PDA is left untreated, it will cause death within the first year of life in 36% of the cases (Stanley, Luis-Fuentes & Darke 2003). In absence of concurrent congenital heart diseases, closure of the ductus can add about ten years to the median lifespan (Saunders et al. 2014).

Histopathological studies showed abnormalities that are believed to cause patency of the ductus in dogs with PDA (Buchanan, Patterson 2003, Buchanan 1978, Gittenberger-De Groot et al. 1985). Normally, the ductus muscle is distributed circumferential and the ductus wall consists of about 98% smooth muscle with a small amount of elastic fibers (Buchanan 2001). In dogs with PDA, hypoplasia and asymmetry of the ductus muscle mass is seen. Furthermore, parts of the ductus muscle wall that should contain smooth muscle cells now have a non-contracting aorta like elastic structure (Buchanan 1978, Gittenberger-De Groot et al. 1985, Buchanan, Patterson 2003). Based on angiographic appearance, the fully patent ductus arteriosus can be classified into four phenotypes (Miller et al. 2006). There is also a phenotypic expression known in which dogs do not have a fully patent ductus, but only a blind diverticulum that communicates with the aorta but not with the pulmonary artery. This is considered to be a 'form fruste' of PDA, representing incomplete closure (Buchanan 2001, Patterson et al. 1971).

Some dog breeds are predisposed for PDA and increased prevalences are reported in certain breeds, including poodles (Patterson et al. 1971) and Chihuahuas (Bomassi et al. 2011). This surely indicates a genetic component in the disease and different modes of inheritance have been suggested and investigated in dogs. In 1968, Patterson et al. performed test matings with affected and non-affected dogs from which results excluded fully penetrant autosomal recessive and sex-linked inheritance but agreed with results that would be expected in autosomal dominant inheritance (Patterson 1968). A second study indicated that hereditary PDA in the dog is not a simple mendelian trait but rather a quasi-continuous or threshold trait with a high degree of heritability (Patterson et al. 1971). In humans, there are a few specific forms of PDA that are caused by mutations in the Myosin Heavy Chain 11 (MYH11) gene (Harakalova et al. 2013). In dogs, no causative gene mutations associated with the occurrence of PDA are described yet.

The Friesian Stabyhoun, a Dutch breed with a relatively small population, seems to have an increased risk of PDA compared to the general dog population. A retrospective study of dogs with PDA at the Utrecht University Clinic for Companion Animals showed a predisposition for PDA in Stabyhoun dogs with an odds ratio of 13,6% (Meijer, Beijerink 2012). A previous study in which pedigree analysis was performed revealed a prevalence of 0,55% and a heritability of 0.4, indicating a possible polygenic inheritance. A genome wide association study of PDA in 24 affected Stabyhouns and 24 control dogs of the same breed resulted in two small chromosomal regions strongly associated with PDA, comprising several candidate genes. A high-throughput sequencing project was performed to detect possible involvement of those genes in PDA, resulting in 157 possible causal single nucleotide polymorphisms (SNPs) with a p-value <0.05 after correction for multiple testing [unpublished personal data].

AIM OF THE STUDY

The aim of this study is to make a start in elucidating the gene mutations that are possibly involved in causing the phenotype of PDA in Friesian Stabyhoun dogs. This will be done by:

1) Validation of 18 single nucleotide polymorphisms in additional Stabyhoun cases and controls
2) Verification of these 18 SNPs in other breeds

MATERIALS AND METHODS

DNA SAMPLES

DNA samples were isolated from 4ml blood with ethylenediaminetetraacetic acid (EDTA), using the Chemagic magnetic separation module I automated nucleic acid isolation system (Perkin Elmer, Groningen, Netherlands) and stored at the department of Clinical Sciences of Companion Animals (Utrecht, Netherlands). The database contained a total of 242 Stabyhoun DNA samples available for genotyping. Forty-eight samples (24 cases and 24 controls) were already genotyped in the previous sequencing project. During this research, DNA of 196 Stabyhouns was genotyped *(Table 1)*.

Table 1: Stabyhoun DNA samples used for genotyping.

Breed	Status	Number
Stabyhoun	Affected	21
Stabyhoun	Control	171
Stabyhoun	Unknown	4
Total samples		**196**

Most of the control samples were collected by large numbers during dog shows. Therefore, we expect this group of samples to give a good representation of the general Dutch Stabyhoun population.

For verification in other breeds, 151 DNA samples from 19 different breeds other than Stabyhouns were available for genotyping, including 77 cases and 73 controls *(Table 2)*. As an extra control, five Stabyhoun samples from which the genotype was already known were also included.

Table 2: Numbers and breeds of DNA samples used for verification.

Breed	Status	N	Breed	Status	N
Bearded Collie	Affected	2	Podenga	Affected	2
	Control	4		Control	1
Border Collie	Affected	3	Poodle	Affected	2
	Control	1		Control	2
Bernese Mountain Dog	Affected	2	Schapendoes	Affected	8
	Control	2		Control	8
Chihuahua	Affected	5	Shetland Sheepdog	Affected	7
	Control	5		Control	7
Cavalier King Charles Spaniel	Affected	7	Saarloos Wolfdog	Affected	2
	Control	7		Control	2
German Brak	Affected	5	Vizsla	Affected	2
	Control	3		Control	2
German Shepherd	Affected	9	West Highland White Terrier	Affected	2
	Control	9		Control	2
Labrador Retriever	Affected	7	Weimaraner dog	Affected	2
	Control	5		Control	2
Polish Lowland Sheepdog	Affected	5	Yorkshire Terrier	Affected	2
	Control	5		Control	2
Newfoundlander	Affected	3			
	Control	4			
	Unknown	1	**Total samples**	**151**	

SINGLE NUCLEOTIDE POLYMORPHISMS

From the 157 possible causal variants found in the previous sequencing project, the top eighteen SNPs with the lowest p-values were screened in both Stabyhoun DNA samples and samples from other breeds *(Table 3)*. To determine on which genes the SNPs were located, the Ensembl Genome Browser was used (Cunningham et al. 2015). To see which effects the mutations could have, the Variant Effect Predictor (VEP) tool was used (*Table 3*) (McLaren et al. 2010, Cunningham et al. 2015). This tool shows the type of variant (e.g. missense variant, downstream gene variant and/or intron variant) and the predicted impact. The SIFT (Sorting Intolerant From Tolerant) score predicts whether the substitution of one amino acid by another amino acid will affect the function of the formed protein. The score ranges from 0 to 1. When the score is <0.05, the amino acid substation is predicted to be damaging and with a score >0.05 it is predicted to be tolerated (Ng, Henikoff 2003).

Table 3: Locations and predicted effects of the top 18 SNPs that were investigated. *The nucleotide changes from each SNP are shown, as well as the variants and the predicted impact they have on the protein function.*

SNP	Chr.	Location	Nucleotide change	Variant	Impact	SIFT score	Prediction
1	5	31280591	G > A	Missense	Moderate	0.1	Tolerated
2	5	31348676	C > T	Missense	Moderate	0	Damaging
3	5	32063394	A > G	Missense	Moderate	0.65	Tolerated
4	5	32116902	T > C	Missense	Moderate	0.04	Damaging
5	5	32330419	G > T	Missense	Moderate	0.03	Damaging
6	5	32358066	A > G	Missense	Moderate	0.12	Tolerated
7	5	32591487	C > T	Missense Missense	Moderate Moderate	0.3 0.08	Tolerated Tolerated
8	5	32778838	G > T	Missense	Moderate	0	Damaging
9	5	32963884	C > T	Missense	Moderate	0.11	Tolerated
10	5	33103525	G > A	Missense	Moderate	0.14	Tolerated
11	5	33241354	G > A	Missense	Moderate	-	-
12	5	34164951	G > T	Missense Missense	Moderate Moderate	0.07 0.09	Tolerated Tolerated
13	5	34185342	C > G	Missense Missense	Moderate Moderate	0.27 0.29	Tolerated Tolerated
14	5	34267512	T > C	Missense	Moderate	0.16	Tolerated
15	5	34289676	A > G	Missense	Moderate	0.27	Tolerated
16	5	34976984	T > C	Missense	Moderate	0.54	Tolerated
17	5	34986321	A > G	Missense	Moderate	0.22	Tolerated
18	19	24467651	C > T	Missense	Moderate	-	-

KASP™ GENOTYPING

KASP™ genotyping, an allele-specific polymerase chain reaction, enables bi-allelic discrimination of SNPs at given loci through competitive binding of two allele-specific forward primers. Both primers are labelled with a different fluorescent dye. After the thermal cycling reaction, an end-point fluorescent read is performed. If the genotype at the given SNP is heterozygous, a mixed fluorescent signal is generated. In case of a homozygous genotype, only one fluorescent signal is generated (LGC Genomics 2014).

Primers that were ordered were based on data from the previous sequencing project. For each SNP, two forward primers and two reverse primers were available. To choose the best working reverse primer for the genotyping, optimization of the primers was performed by testing the protocol on a selection of sixteen Stabyhoun samples (8 cases and 8 controls) that were already genotyped in the sequencing project, thereby making it possible to compare and confirm the results from the optimization.

DNA was diluted 15 times by adding Milli-Q® water. Primermixes contained the two forward (F) primers A1 and A2, one reverse (R) primer (C1 or C2 *(Table 4)*) and Milli-Q® water. Polymerase chain reactions were performed on the GeneAmp® 9700 PCR machine (Life Technologies, Bleiswijk, Netherlands) in a 4 µl volume reaction containing KASPar mix (2x) (LGC Genomics, Middlesex, UK), primermix (IDT Integrated DNA Technologies, Coralville (Iowa), USA) and genomic DNA (1:15). For negative control, extra Milli-Q® water monsters were added.
The following cycling protocol was used for the PCR: 15 minutes at 94°C, followed by 10 cycles of 20 seconds at 94°C and 60 seconds at 61°C with a -0,6°C drop per cycle, followed by 30 cycles of 10 seconds at 94°C and 60 seconds at 55°C and ending with 10 cycles of 20 seconds at 94°C and 60 seconds at 57°C. The reaction was maintained on 20°C. In case of too many false positive results, a shortened version of the protocol mentioned above was used, skipping the last 10 cycles and ending after 25 instead of 30 cycles of the middle part of the protocol (see *Appendix 1* for complete protocols).
The end-point read was performed with the CFX384™ Real-Time System (Bio-Rad, Veenendaal, the Netherlands) and results were analysed with CFX Manager™ Software 3.1 (Bio-Rad).

Table 4: Forward (F) and reverse (R) primers used for KASP™ genotyping.
The reverse primers that were chosen after optimization are bold.

SNP	Primer name	F/R	Sequence
1	5_31280591_A1	F	GAAGGTGACCAAGTTCATGCTAAGTGGAGATGGTGGCCTCATC
	5_31280591_A2	F	GAAGGTCGGAGTCAACGGATTAAAGTGGAGATGGTGGCCTCATT
	5_31280591_C1	**R**	**GGACCTCAGGAAGCACGGGTT**
	5_31280591_C2	R	TCAGCAGGAGGGACCTCAGGAA
2	5_31348676_A1	F	GAAGGTGACCAAGTTCATGCTCATGCCCATGGTGTCCTGCC
	5_31348676_A2	F	GAAGGTCGGAGTCAACGGATTCCATGCCCATGGTGTCCTGCT
	5_31348676_C1	**R**	**GCGCCAAGCAGCGGCGGAT**
	5_31348676_C2	R	AGCGGCGGATGCTCTCCACAT
3	5_32063394_A1	F	GAAGGTGACCAAGTTCATGCTCCTGCCAGCCTCCCAGCA
	5_32063394_A2	F	GAAGGTCGGAGTCAACGGATTCCTGCCAGCCTCCCAGCG
	5_32063394_C1	R	TGGGGAAAGCAGAGCAACTTGAGTA
	5_32063394_C2	**R**	**CAACTTGAGTAGTGGGGAGCAGTT**
4	5_32116902_A1	F	GAAGGTGACCAAGTTCATGCTGCGCTCTCCAGCTGGCAGTA
	5_32116902_A2	F	GAAGGTCGGAGTCAACGGATTCGCTCTCCAGCTGGCAGTG
	5_32116902_C1	R	ACGAGGGCAGCTGCTACTGGTT
	5_32116902_C2	**R**	**GACCTGGGAGGAGGCCGAGAA**
5	5_32330419_A1	F	GAAGGTGACCAAGTTCATGCTACCTGCTCCCGCGCCCG
	5_32330419_A2	F	GAAGGTCGGAGTCAACGGATTACCTGCTCCCGCGCCCT
	5_32330419_C1	**R**	**GGGGCCGGGCGAGGGGAA**

	5_32330419_C2	R	GGGCGAGGGGAAGAGAGAGAA
6	5_32358066_A1	F	GAAGGTGACCAAGTTCATGCTCCATCACCACCCTTGCAGCCA
	5_32358066_A2	F	GAAGGTCGGAGTCAACGGATTCATCACCACCCTTGCAGCCG
	5_32358066_C1	**R**	**CTCCCATTCTTCAGCCCCCACT**
	5_32358066_C2	R	CCGGACGTGCCTCCCATTCTT
7	5_32591487_A1	F	GAAGGTGACCAAGTTCATGCTCCAGGCCGTGGGCCG
	5_32591487_A2	F	GAAGGTCGGAGTCAACGGATTGCTCCAGGCCGTGGGCCA
	5_32591487_C1	**R**	**TGTCTGAAATGCCCATGGAGAGAGA**
	5_32591487_C2	R	GAAGCCTGTGTCTGAAATGCCCAT
8	5_32778838_A1	F	GAAGGTGACCAAGTTCATGCTAGACAGTTCAATCTGATTCCTCCTG
	5_32778838_A2	F	GAAGGTCGGAGTCAACGGATTCAGACAGTTCAATCTGATTCCTCCTT
	5_32778838_C1	**R**	**GATGGTGCTCAGCTTGCTCACACA**
	5_32778838_C2	R	CTCTGAAACAGATGGTGCTCAGCTT
9	5_32963884_A1	F	GAAGGTGACCAAGTTCATGCTGAGTCCAGAGACAGAGCCGG
	5_32963884_A2	F	GAAGGTCGGAGTCAACGGATTAGAGTCCAGAGACAGAGCCGA
	5_32963884_C1	R	CAGGGCACCCCTGAATGAGGAT
	5_32963884_C2	**R**	**GTTCACTCCCCCGGCCCCTA**
10	5_33103525_A1	F	GAAGGTGACCAAGTTCATGCTCCAGGAGGCCACTGGCTC
	5_33103525_A2	F	GAAGGTCGGAGTCAACGGATTGTCCAGGAGGCCACTGGCTT
	5_33103525_C1	R	GGCCGCCTGGACCTCGCAA
	5_33103525_C2	**R**	**GCAACACACTGTGGCAGGAGCTT**
11	5_33241354_A1	F	GAAGGTGACCAAGTTCATGCTACGGAGAGTAAGACCAGTGACG
	5_33241354_A2	F	GAAGGTCGGAGTCAACGGATTCACGGAGAGTAAGACCAGTGACA
	5_33241354_C1	R	CTACTCCGACTGCGGCGGCT
	5_33241354_C2	**R**	**CTGCGGCGGCTGTGTCTCGTT**
12	5_34164951_A1	F	GAAGGTGACCAAGTTCATGCTATGAAGTTATACATACCTGCTGGCC
	5_34164951_A2	F	GAAGGTCGGAGTCAACGGATTGATGAAGTTATACATACCTGCTGGCA
	5_34164951_C1	**R**	**CAAGAATAAGGTGCTAATCCTCATCTGTA**
	5_34164951_C2	R	GTAACTTGGTGGGGTCAGGGCAA
13	5_34185342_A1	F	GAAGGTGACCAAGTTCATGCTGGCCACTTCCCTTGCAGTTTG
	5_34185342_A2	F	GAAGGTCGGAGTCAACGGATTGGCCACTTCCCTTGCAGTTTC
	5_34185342_C1	R	ATGAGAGACGAGCCCGAGCTGA
	5_34185342_C2	**R**	**GCCCGAGCTGAGGGCCAGAA**
14	5_34267512_A1	F	GAAGGTGACCAAGTTCATGCTCCCATGGACACCCCTGCCT
	5_34267512_A2	F	GAAGGTCGGAGTCAACGGATTCCATGGACACCCCTGCCC
	5_34267512_C1	R	TGGAGGGGGGCAGAGGCCT
	5_34267512_C2	**R**	**GGCCGGCGGAGAGCGCTT**
15	5_34289676_A1	F	GAAGGTGACCAAGTTCATGCTGACATACCTCGGACACATAGGAT
	5_34289676_A2	F	GAAGGTCGGAGTCAACGGATTGACATACCTCGGACACATAGGAC
	5_34289676_C1	**R**	**CCAGGAGGCCCAACAGTGACAA**
	5_34289676_C2	R	GCCCAACAGTGACAAAGAGTGGAT
16	5_34976984_A1	F	GAAGGTGACCAAGTTCATGCTCCAAGCTGAAGCACACAGCA
	5_34976984_A2	F	GAAGGTCGGAGTCAACGGATTCCAAGCTGAAGCACACAGCG
	5_34976984_C1	**R**	**GGGGTGCCATCGAACACTGGAA**
	5_34976984_C2	R	CTGGAATACAGAGGGCCGCCTA
17	5_34986321_A1	F	GAAGGTGACCAAGTTCATGCTCAAGAGATAAATGGGAGGAGAGTGAT
	5_34986321_A2	F	GAAGGTCGGAGTCAACGGATTAAGAGATAAATGGGAGGAGAGTGAC
	5_34986321_C1	**R**	**GCTGGTGGAAGAATTCAGTTGGCTA**
	5_34986321_C2	R	AAGAATTCAGTTGGCTACTGCTGTTGTA
18	19_24467651_A1	F	GAAGGTGACCAAGTTCATGCTGTGTGTCAGCAACTCCATCCG
	19_24467651_A2	F	GAAGGTCGGAGTCAACGGATTGGTGTGTCAGCAACTCCATCCA
	19_24467651_C1	R	GGATGGCCAAACCCACATGCTATTT
	19_24467651_C2	**R**	**CAAACCCACATGCTATTTGGCTCTGAT**

STATISTICS

To compare allele and genotype results between cases and controls, a Chi square test was performed. In case of less than five observations, a Fisher's exact test was used. In both tests, a p value <0.05 was considered to be statistically significant. Correction for multiple testing was performed with use of the Bonferroni method.

RESULTS

Eventually, all SNPs except number 5 were screened by KASP™ genotyping. For SNP 2, 9 and 13, the shortened version of the PCR protocol was used. For SNP 5, the KASP™ genotyping method did not work. Therefore we tried to genotype the samples with Sanger sequencing, but because of some difficulties we have no data yet.

Figure 2 shows an example of the end-point read images that were obtained at analysis with the CFX Manager™ Software 3.1. The three coloured clusters represent the three different genotypes: homozygous wild type (yellow), heterozygous (green) and homozygous mutant (blue). Based on these images, the genotype from every single sample was determined for each of the SNPs. Allele frequency and genotype frequency were calculated on the basis of the total lists of genotypes per SNP.

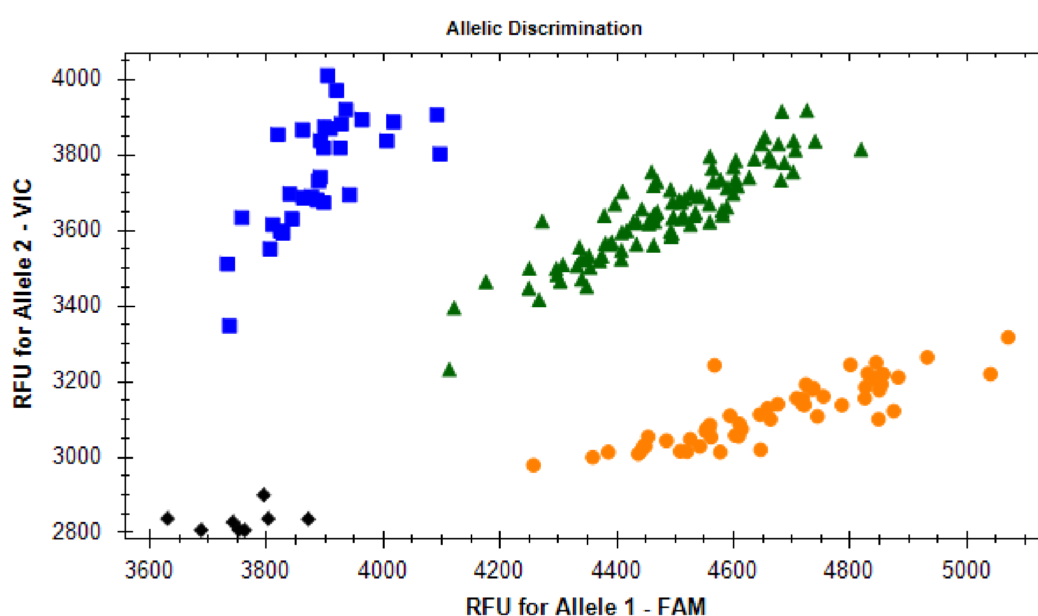

Figure 2: Example of an end-point read image from KASP™ genotyping. *The yellow dots represent the homozygous wild type samples, the green triangles the heterozygous samples and the blue squares the homozygous mutant samples. The black squares in the bottom left corner represent the negative control samples.*

VALIDATION IN STABYHOUNS

In total, 196 Stabyhouns were genotyped. One case and one control sample were already genotyped in the previous sequencing project and served as an additional control. The remaining 194 samples have not been genotyped before.

The statistical analysis is based on results from 21 cases and 171 controls. The status of four samples was unknown, therefore these samples were excluded from statistical analysis. *Table 5* shows the genes on which the SNPs are located, just as allele frequencies and p-values. A significant difference between PDA cases and controls in allele frequency was shown for SNP 10 (p=0.041) and SNP 14 (p=0.045), but only before correction for multiple testing. In genotype

frequencies, no significances were demonstrated. *Table 6* shows the allele frequencies from the significant SNPs on which the p-values are based.

VERIFICATION IN OTHER BREEDS

All SNPs were screened in 151 samples from 19 different breeds other than Stabyhouns and five Stabyhoun samples as extra control. Statistical analysis is based on results from 150 samples, including 77 cases and 73 controls. The five Stabyhoun samples were excluded from statistical analysis, as well as one New Foundlander sample which had an unknown status. A significant difference in allele frequency between PDA cases and controls was shown for SNP 9 (p=0.018) and a significant difference in genotype frequency was shown for SNP 16 (p=0.023). Both were only significant before correction for multiple testing *(Table 5)*. Allele frequencies and genotype frequencies from significant SNPs are shown in *Table 6* and *Table 7*.

Table 5: Gene locations and results from statistical analysis from validation in Stabyhouns and verification in other breeds. *Allelic (P-ALL) and genotypic (P-GEN) p-values before and after Bonferroni correction (BC) for multiple testing are shown.*
** = statistical analysis with Fisher's exact test. NS = not significant.*

		Stabyhouns				Other breeds			
SNP	Gene	P-ALL	BC	P-GEN	BC	P-ALL	BC	P-GEN	BC
1	NLRP1	NS	NS	NS*	NS	NS	NS	NS*	NS
2	DHX33	NS	NS	NS*	NS	NS	NS	NS*	NS
3	SLC16A11	NS	NS	NS*	NS	NS*	NS	NS*	NS
4	ASGR2	NS	NS	NS	NS	NS	NS	NS	NS
5	TMEM256-PLSCR3	No data		No data		No data		No data	
6	C17orf74	NS	NS	NS*	NS	NS	NS	NS*	NS
7	EFNB3	NS	NS	NS*	NS	NS	NS	NS*	NS
8	KCNAB3	NS	NS	NS*	NS	NS	NS	NS	NS
9	PER1	NS	NS	NS*	NS	**1,8E-02***	NS	NS	NS
10	ARHGEF15	**4,1E-02**	NS	NS*	NS	NS*	NS	NS	NS
11	MYH10	NS	NS	NS*	NS	NS	NS	NS	NS
12	USP43	NS	NS	NS*	NS	NS	NS	NS	NS
13	USP43	NS	NS	NS*	NS	NS	NS	NS	NS
14	GLP2R	**4,5E-02**	NS	NS*	NS	NS	NS	NS*	NS
15	GLP2R	NS	NS	NS*	NS	NS	NS	NS*	NS
16	ADPRM	NS	NS	NS*	NS	NS	NS	**2,3E-02**	NS
17	TMEM220	NS	NS	NS*	NS	NS	NS	NS	NS
18	GSTK1	NS	NS	NS	NS	NS	NS	NS	NS

NLRP1= NLR family, pyrin domain containing 1, DHX33= DEAH (Asp-Glu-Ala-His) Box Polypeptide 33, SLC16A11= solute carrier family 16, member 11, ASGR2= asialoglycoprotein receptor 2, C17orf74= chromosome 17 open reading frame 74, EFNB3= ephrin-B3, KCNAB3= K(+) Channel Subunit Beta-3, PER1= period circadian clock 1, ARHGEF15= Rho Guanine Nucleotide Exchange Factor 15, MYH10= Myosin Heavy Chain 10, USP43= ubiquitin specific peptidase 43, GLP2R= glucagon-like peptide 2 receptor, ADPRM= ADPRibase-Mn, TMEM220= transmembrane protein 220, GSTK1= glutathione S-transferase kappa 1.

Table 6: Allele frequencies from significant SNPs in Stabyhouns and other breeds.
AF_{ref} = allele frequency of wild type alleles, AF_{mut} = allele frequency of mutant alleles

Stabyhouns					**Other breeds**				
	Controls		Cases			Controls		Cases	
SNP	AF_{wild}	AF_{mut}	AF_{wild}	AF_{mut}	SNP	AF_{wild}	AF_{mut}	AF_{wild}	AF_{mut}
10	0,53	0,47	0,70	0,30	9	0,93	0,07	0,99	0,01
14	0,55	0,45	0,71	0,29					

Table 7: Genotype frequencies from the significant SNP in other breeds.
GF_{wild} = homozygous wild type, GF_{het} = heterozygous, GF_{mut} = homozygous mutant

SNP 16	**Controls**	**Cases**
GF_{wild}	0,63	0,53
GF_{het}	0,16	0,36
GF_{mut}	0,21	0,11

DISCUSSION

Patent ductus arteriosus is thought to be a complex polygenic disease. When the causative mutations can be determined, a genetic test can be developed to screen Friesian Stabyhouns and possibly other (predisposed) dog breeds. Such a test is very useful because patients and carriers can be detected and excluded from breeding, so that the disease will not be passed on to the offspring.

In previous GWAS and sequencing studies, 157 possible causative single nucleotide polymorphisms were found. SNPs are point mutations that can occur in protein coding regions as well as in non-protein coding regions of genes. Mutations occurring in protein coding regions (exons) can directly influence the proteins that are formed. When a single nucleotide change results in a codon that codes for a different amino acid, it is called a missense mutation. When a single nucleotide change results in a change to a premature stop codon, it is called a nonsense mutation (Russell 2009). Insertion of different amino acids in a polypeptide can possibly impair normal functioning of the polypeptide, thereby leading to abnormalities in the functioning of cells and body tissues.

During this research, the top 18 possible causative SNPs with the lowest p-values after correction for multiple testing were screened in DNA from Stabyhouns and other breeds in order to detect which variants they carried. The objective was to find out if one or more of these 18 SNPs could be involved in causing the phenotype of PDA in the Friesian Stabyhoun. All but the last SNP are located on chromosome 5, within a region of about 3 Mb. The last SNP is located on chromosome 19. Because the location and nucleotide change of all SNPs was known, it was possible to investigate amino acid changes and involved genes. All SNPs turned out to be missense mutations, occurring in protein coding regions *(Table 3)*.

VALIDATION IN STABYHOUNS

In the validation in additional Stabyhoun cases and controls, a significant difference between cases and controls in allele frequency was seen for SNP 10 and SNP 14.

SNP 10 is characterized by a nucleotide change from guanine (G) to adenine (A). In the amino acid sequence, this results in a change from arginine (Arg) to glutamine (Gln). There is a slight difference between the characteristics of these two amino acids. Both are polar, but arginine is charged and basic whereas glutamine is acidic (Betts, Russell 2003). The SNP is located on the Rho Guanine Nucleotide Exchange Factor (GEF) 15 gene (ARHGEF15). This gene encodes for a specific guanine nucleotide exchange factor for RhoA, a Rho GTPase that works through G-protein coupled receptors. ARHGEF15 is essentially expressed in vascular smooth muscle cells (Loirand et al. 2008, Kather, Kroll 2013), for example in the vascular smooth muscle of the coronary artery, and is involved in the regulation of vascular smooth muscle contractility (Weizmann Institute of Science 2015, Ogita et al. 2003). In vascular smooth muscle cells, it interacts with ephrin A4 (EFNA4). EFNA4 is a member of the family of receptor tyrosine kinases, which are very important for migration, repulsion and adhesion during neuronal, vascular and epithelial development (Weizmann Institute of Science 2015). It also plays a role in the contractility of vascular smooth muscles (Ogita et al. 2003). It is expressed in vascular smooth muscle cells in many organs, including the fetal heart (Weizmann Institute of Science 2015, Ogita et al. 2003).

In SNP 14, thymine (T) is replaced by cytosine (C). Due to this change, instead of cysteine (Cys) the amino acid arginine (Arg) is inserted. These amino acids are quite different from each other. Cysteine is small, polar, hydrophobic and contains sulphur. Arginine is, as mentioned above, polar, charged and basic (Betts, Russell 2003). This SNP is located on the Glucagon-Like Peptide 2 Receptor (GLP2R) gene. GLP2R is the receptor for GLP2, a peptide that is produced by intestinal endocrine cells and that maintains the growth and absorptive function of the intestinal mucosal villus epithelium. It is widely expressed in several organs, including heart, kidney, liver and smooth and skeletal muscles (Weizmann Institute of Science 2015).

Although the Chi square test showed significance in two SNPs for allele frequencies, the p-values were very high and only just below 0.05. Obviously, they did not remain significant after correction for multiple testing. On the other hand, the data from the sequencing project showed very low p-values after correction for multiple testing for all eighteen SNPs. This included significance in both allele and genotype frequency, ranging from 1.2E-05 to 6.5E-06 for allele frequency and from 1.0E-04 to 8.8E-05 for genotype frequency [unpublished personal data]. When we combined the data from the sequencing project with our own data, we see much lower p-values for almost all SNPs, even after correction for multiple testing *(Table 8)*. In combined data, statistical analysis is based on 44 cases and 194 controls. Allelic p-values range from 2.7E-02 to 8.6E-03, with one SNP being significant only before correction for multiple testing and one SNP being not significant at all. Looking at the genotypes, p-values range from 4.8E-02 to 4.1E-03, with six SNPs being significant only before correction for multiple testing and one SNP not being significant at all. One explanation for the big differences in p-values could be the rate between the number of cases and controls. Analysis of the data from the sequence project was based on a balanced number of 24 cases and 24 controls. On the other hand, the analysis of the data of this study is based on only 21 cases together with the large number of 171 controls, which is eight times as much controls as cases. This imbalance has a negative effect on the p-values, decreasing the significance. Another reason could be that the samples from the sequencing project were not randomly selected and that the population from which the samples in this study were used is different than the population from which samples for the sequencing project were selected.

Because all three genotypes (homozygous wild type, heterozygous and homozygous mutant) are seen in both cases and controls, the cause of PDA cannot be explained by just the significant SNPs as the only causative mutation(s). However, we would expect to see the mutant alleles more than the wild type alleles in the cases. Looking at the allele frequencies in both significant SNPs, it is remarkable that the mutant allele is less present than the wild type allele *(Table 6)*. It could be that these mutations are linked to other mutations involved in the cause of PDA and are in some way protective against the development of PDA. In this way, we would expect to see less mutant alleles in PDA cases because the protective mutation did not occur. The effect of the mutations is predicted to be tolerated (*Table 3*), which suits the suggestion of being protective.

Looking at the genes in which the SNPs are located, at the start of the research we expected SNPs number 10 and 11 to be most interesting, because of their location on genes involved in cardiac vascular function. Especially SNP number 11 seemed interesting, because of known human mutations in a similar gene. SNP number 11 is located on the non-muscle Myosin Heavy Chain 10 (MYH10) gene, a gene that belongs to the myosin superfamily. Myosins are actin-dependent motor proteins involved in muscle contraction. Mutations in the MYH10 gene are associated with developmental defects in heart and brain (Weizmann Institute of Science 2015, Ma, Adelstein 2014). Knockout studies in mice showed that MYH10 is essential for cardiac and brain development and a point mutation in the gene is shown to cause major heart defects in

mice, including ventricular septal defect and a double outlet of the right ventricle (Ma, Adelstein 2014). In humans, mutations in another member of the myosin family, the myosin heavy chain 11 (MYH11) gene, are known to cause specific forms of PDA (Harakalova et al. 2013). However, in this research, no significance between cases and controls was shown in both the Stabyhouns and other breeds for SNP 11.

For now, in spite of the minimal significance found, especially the SNP on the ARHGEF15 gene seems interesting because of its expression and function in vascular smooth muscle cells. A direct or indirect (via interaction with EFNA4) dysfunction of this gene could lead to an impaired development and contractility of vascular smooth muscles, which is also the case in PDA. For further research, it would be informative to screen the SNPs in a large group of Stabyhouns with a balanced number of cases and controls, to see what significant differences are found then. For example, looking at the predicted effects, the amino acid substitution of some SNPs that did not appear to be significant are predicted to be damaging *(Table 3)*. In the combined data, these SNPs do appear to be significant, so it would be interesting to see if any significant differences can be detected in a more balanced group of Stabyhouns. Also, in this research, only 18 of 157 possible variants were screened. It's not ruled out that other variants that were not screened here are involved, so screening more variants could also be something for further projects. In case of significant correlations, expression of involved genes in vascular smooth muscle can be measured by use of qPCR. With this method, there can also be looked at differences in expression levels from affected and unaffected dogs. Eventually, knock-out studies in mice can give information about the function of certain genes in the phenotype of PDA.

Table 8: Results from statistical analysis of the data from this study combined with the data from the sequencing project. *Allelic (P-ALL) and genotypic (P-GEN) p-values before and after Bonferroni correction (BC) for multiple testing.*
** = statistical analysis with Fisher's exact test. NS = not significant.*

Name	Gene	P-ALL	BC	P-GEN	BC
SNP 1	NLRP1	3,5E-04	6,2E-03	1,9E-03	3,3E-02
SNP 2	DHX33	3,2E-04	5,7E-03	2,0E-03*	3,6E-02
SNP 3	SLC16A11	9,7E-05	1,7E-03	2,7E-04	4,8E-03
SNP 4	ASGR2	8,0E-05	1,4E-03	2,3E-04	4,1E-03
SNP 5	TMEM256-PLSCR3	No data	No data	No data	No data
SNP 6	C17orf74	2,7E-04	4,8E-03	1,1E-03	2,0E-02
SNP 7	EFNB3	1,8E-04	3,2E-03	1,3E-03*	2,3E-02
SNP 8	KCNAB3	4,8E-04	8,6E-03	3,7E-03*	NS
SNP 9	PER1	9,0E-04	1,6E-02	2,8E-03	NS
SNP 10	ARHGEF15	7,8E-05	1,4E-03	6,9E-04*	1,2E-02
SNP 11	MYH10	1,5E-03	2,7E-02	4,5E-03	NS
SNP 12	USP43	4,4E-04	7,9E-03	3,3E-03*	NS
SNP 13	USP43	3,4E-04	6,2E-03	1,9E-03*	3,4E-02
SNP 14	GLP2R	3,3E-04	6,0E-03	2,2E-03*	3,9E-02
SNP 15	GLP2R	3,7E-04	6,6E-03	1,9E-03*	3,5E-02
SNP 16	ADPRM	6,0E-03	NS	1,2E-02	NS
SNP 17	TMEM220	1,6E-03	3,0E-02	4,5E-03	NS
SNP 18	GSTK1	NS	NS	NS	NS

VERIFICATION IN OTHER BREEDS

To see if the SNPs found in the Stabyhouns could also be important in PDA in other breeds, cases and controls from other breeds were also genotyped in order to see which variants they carried. A significance between cases and controls in allele frequency was seen for SNP 9 and a significant difference in genotype frequency was seen for SNP 16.

In SNP 9, there is a nucleotide change from cysteine (C) to thymine (T). In amino acid sequence, proline (Pro) changes to serine (Ser). Proline is a small amino acid, whereas serine is tiny, polar and hydroxylic (Betts, Russell 2003). The SNP is located on the Period Circadian Clock 1 (PER1) gene, a member of the Period family. It is the primary circadian pacemaker in the mammalian brain. With the highest level in skeletal muscle, the gene is widely expressed in several organs, including heart, brain, lung, liver, placenta, pancreas, kidney, spleen, small intestine, prostate, testis and ovary (Weizmann Institute of Science 2015).

The nucleotide change in SNP 16 is similar to that from SNP 14, namely thymine (T) changes to cytosine (C). This also results in the same amino acid change as in SNP 14; cysteine changes to arginine. It is located on the ADP-Ribose/CDP-Alcohol Diphosphatase, Manganese-Dependent (ADPRM) gene, which is a hydrolyser of several non-reducing ADP-sugars and CDP-glucose. Tissue expression sites include the heart (Weizmann Institute of Science 2015).

Looking at the results from statistical analysis, there are a few things that can be discussed. First, there is the significance in allele frequency for SNP number 9. In the control group, 136 wild type alleles (C) and 10 mutant type alleles (T) were found. In the cases, 150 wild type alleles and 2 mutant type alleles were found. The difference between the groups is very small. The size of both groups is almost the same so that cannot have a big influence. Even though the result of the Fisher's exact test was significant, the difference in amount of wild type and mutant alleles is too small to draw any conclusion. Thus, based on the results of this study, it is not yet clear if this SNP could actually play a role in the cause of PDA in other breeds.

Second, we have the significance in genotype frequency for SNP number 16. Looking at the genotype frequencies to see which genotype makes the significant difference between cases and controls, both the homozygous wild type and the homozygous mutant genotypes do not differ that much *(Table 7)*. This suggests that these genotypes both are 'healthy'. The biggest difference is seen between the percentages heterozygotes. However, it would be very unlikely that the heterozygous genotype carries the phenotype of PDA, because it consists of two alleles that are expected to be healthy. Most likely, the results are false positive and no conclusion can be drawn.

The SNPs that were found significant in Stabyhouns did not appear to be significant in the other breeds. The mutant allele from SNP number 10 is practically not seen in other breeds. Only two control samples from two different breeds were heterozygous and showed one mutant allele. All other 149 samples, including all cases, were homozygous wild type. Because the mutant allele is hardly seen in both cases and controls, this results exclude SNP 10 from playing a role in PDA in other breeds.

Because the analysis is based on just cases and controls without distinction between different dog breeds, the results do only give information about a group of other breeds than Stabyhouns in general and not about any particular breed. Because there were just a few samples per breed available *(Table 2)*, it was not possible to test for significant differences per breed. When variants that are significant in Stabyhouns are screened in other breeds, it would give more information if bigger groups of different breeds were screened. In that way, it can be tested if certain variants could play a role in specific breeds in the phenotype of PDA.

CONCLUSION

In this research, a start was made in elucidating the gene mutations that are involved in the development of PDA in the Friesian Stabyhoun. The Stabyhoun dog might be a good model for other dog breeds and PDA in human medicine, because of the small genetic base and the high prevalence of PDA in this breed. From the top 18 SNPs that were screened, two SNPs showed a small significant difference between Stabyhoun cases and controls in allele frequency. These results show possible involvement of these gene mutations in the development of PDA. However, PDA is a complex polygenic disease, so further research is needed to find out the exact mechanism of the genetic cause of this disease. More SNPs can be screened, expression of involved genes can be measured with qPCR and knock-out studies can provide information about the effect of certain gene mutations. The ultimate goal is to develop a genetic test to screen Stabyhoun dogs, thereby making it possible to exclude patients and carriers from breeding.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Frank van Steenbeek for all the time and support during my research project, giving me the chance to learn so many different techniques. Also a special thanks for Manon Vos-Loohuis for helping me with all the lab work. Thanks to all people at the lab who made me feel very welcome and gave me the opportunity to work at the lab. It was a great and educational experience.

REFERENCES

Betts, M.J. & Russell, R.B. 2003, "Amino acid properties and consequences of substitutions" in *Bioinformatics for geneticists*, eds. M.J. Betts & R.B. Russell, Wiley, Chichester, Eng.; Hoboken, N.J., pp. 311-342.

Bökenkamp, R., Deruiter, M.C., Van Munsteren, C. & Gittenberger-De Groot, A.C. 2010, "Insights into the pathogenesis and genetic background of patency of the ductus arteriosus", *Neonatology,* vol. 98, no. 1, pp. 6-17.

Bomassi, E., Libermann, S., Bille, C. & Rattez, E. 2011, "Patent ductus arteriosus in a family of Chihuahuas", *Journal of Small Animal Practice,* vol. 52, no. 4, pp. 213-219.

Buchanan, J. 1999, "Prevalence of cardiovascular disorders. In: Canine and feline cardiology. " in *Canine and feline cardiology* WB Saunders, Philadelphia PA, pp. 457-470.

Buchanan, J.W. 2001, "Patent Ductus Arteriousus Morphology, Pathogenesis, Types and Treatment", *Journal of Veterinary Cardiology,* vol. 3, no. 1, pp. 7-16.

Buchanan, J.W. 1978, "Morphology of the ductus arteriosus in fetal and neonatal dogs genetically predisposed to patent ductus arteriosus", *Birth Defects: Original Article Series,* vol. 14, no. 7, pp. 349-360.

Buchanan, J.W. & Patterson, D.F. 2003, "Etiology of Patent Ductus Arteriosus in Dogs", *Journal of Veterinary Internal Medicine,* vol. 17, no. 2, pp. 167-171.

Clyman, R.I. 2006, "Mechanisms regulating the ductus arteriosus", *Biology of the neonate,* vol. 89, no. 4, pp. 330-335.

Coceani, F. & Baragatti, B. 2012, "Mechanisms for Ductus Arteriosus Closure", *Seminars in perinatology,* vol. 36, no. 2, pp. 92-97.

Cunningham et al., F. 2015, "Ensembl 2015.", *Nucleic Acids Research,* vol. 43, pp. D662-D669.

Dzialowski, E.M., Sirsat, T., van der Sterren, S. & Villamor, E. 2011, "Prenatal cardiovascular shunts in amniotic vertebrates", *Respiratory Physiology and Neurobiology,* vol. 178, no. 1, pp. 66-74.

Gittenberger-De Groot, A.C., Strengers, J.L.M., Mentink, M., Poelmann, R.E. & Patterson, D.F. 1985, "Histologic studies on normal and persistent ductus arteriosus in the dog", *Journal of the American College of Cardiology,* vol. 6, no. 2, pp. 394-404.

Harakalova, M., Van Der Smagt, J., De Kovel, C.G.F., Slot, R.V., Poot, M., Nijman, I.J., Medic, J., Joziasse, I., Deckers, J., Roos-Hesselink, J.W., Wessels, M.W., Baars, H.F., Weiss, M.M., Pals, G., Golmard, L., Jeunemaitre, X., Lindhout, D., Cuppen, E. & Baas, A.F. 2013, "Incomplete segregation of MYH11 variants with thoracic aortic aneurysms and dissections and patent ductus arteriosus", *European Journal of Human Genetics,* vol. 21, no. 5, pp. 487-493.

Kather, J.N. & Kroll, J. 2013, "Rho guanine exchange factors in blood vessels: Fine-tuners of angiogenesis and vascular function", *Experimental cell research,* vol. 319, no. 9, pp. 1289-1297.

Kittleson, Mark D., Kienle,Richard D., 1998, "Patent ductus arteriosus. In: Small animal cardiovascular medicine" in *Small animal cardiovascular medicine* Mosby, St. Louis, MO, pp. 218-230.

LGC Genomics 2014, , *How does KASP work*. Available: http://www.lgcgroup.com/LGCGroup/media/PDFs/Products/Genotyping/kasp-explanation-fact-sheet.pdf?ext=.pdf [Downloaded at April 7, 2015].

Loirand, G., Scalbert, E., Bril, A. & Pacaud, P. 2008, "Rho exchange factors in the cardiovascular system", *Current Opinion in Pharmacology,* vol. 8, no. 2, pp. 174-180.

Ma, X. & Adelstein, R.S. 2014, "A point mutation in myh10 causes major defects in heart development and body wall closure", *Circulation: Cardiovascular Genetics,* vol. 7, no. 3, pp. 257-265.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P. & Cunningham, F. 2010, "Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor", *Bioinformatics,* vol. 26, no. 16, pp. 2069-2070.

Meijer, M. & Beijerink, N. 2012, "Persisterende ductus arteriosus bij de hond: een retrospectieve studie naar klinisiche presentatie, diagnostiek, en vergelijking van toegepaste therapeutische mogelijkheden bij 102 honden.", *Netherlands journal of veterinary science,* , pp. 376-383.

Miller, M.W., Gordon, S.G., Saunders, A.B., Arsenault, W.G., Meurs, K.M., Lehmkuhl, L.B., Bonagura, J.D. & Fox, P.R. 2006, "Angiographic classification of patent ductus arteriosus morphology in the dog", *Journal of Veterinary Cardiology,* vol. 8, no. 2, pp. 109-114.

Ng, P.C. & Henikoff, S. 2003, "SIFT: Predicting amino acid changes that affect protein function", *Nucleic acids research,* vol. 31, no. 13, pp. 3812-3814.

Ogita, H., Kunimoto, S., Kamioka, Y., Sawa, H., Masuda, M. & Mochizuki, N. 2003, "EphA4-mediated Rho activation via Vsm-RhoGEF expressed specifically in vascular smooth muscle cells", *Circulation research,* vol. 93, no. 1, pp. 23-31.

Oliveira, P., Domenech, O., Silva, J., Vannini, S., Bussadori, R. & Bussadori, C. 2011, "Retrospective review of congenital heart disease in 976 dogs", *Journal of Veterinary Internal Medicine,* vol. 25, no. 3, pp. 477-483.

Patterson, D.F. 1968, "Epidemiologic and genetic studies of congenital heart disease in the dog.", *Circulation research,* vol. 23, no. 2, pp. 171-202.

Patterson, D.F., Pyle, R.L., Buchanan, J.W., Trautvetter, E. & Abt, D.A. 1971, "Hereditary patent ductus arteriosus and its sequelae in the dog.", *Circulation research,* vol. 29, no. 1, pp. 1-13.

Russell, P.J. 2009, "DNA mutation" in *iGenetics : a molecular approac* Pearson Education, Upper Saddle River, N.J.; Harlow, pp. 131-145.

Saunders, A.B., Gordon, S.G., Boggess, M.M. & Miller, M.W. 2014, "Long-term outcome in dogs with patent ductus arteriosus: 520 Cases (1994-2009)", *Journal of Veterinary Internal Medicine,* vol. 28, no. 2, pp. 401-410.

Schneider, D.J. 2012, "The Patent Ductus Arteriosus in Term Infants, Children, and Adults", *Seminars in perinatology,* vol. 36, no. 2, pp. 146-153.

Stanley, B.J., Luis-Fuentes, V. & Darke, P.G.G. 2003, "Comparison of the Incidence of Residual Shunting Between Two Surgical Techniques Used for Ligation of Patent Ductus Arteriosus in the Dog", *Veterinary Surgery,* vol. 32, no. 3, pp. 231-237.

Weizmann Institute of Science 2015, , *GeneCards*. Available: http://www.genecards.org/.

APPENDIX 1: PROTOCOL KASP™ GENOTYPING

Primermix 50 µL:

- 6 µL forward primer A1 (100 µM)
- 6 µL forward primer A2 (100 µM)
- 15 µL reverse primer C1 or C2 (100 µM)
- 23 µL Milli-Q

PCR reaction volume per well:

KASPar mix (2x)	2 µL
Primermix	0,0055 µL
gDNA (1:15)	2 µL

Cycle protocol:

94ºC 30 minutes

94 ºC 20 seconds
61ºC 60 seconds
Drop -0,6 ºC per cycle
] 10 cycles

94 ºC 10 seconds
55ºC 60 seconds
] 30 cycles

94 ºC 20 seconds
57ºC 60 seconds
] 10 cycles

20ºC ∞

Shortened protocol:

94ºC 30 minutes

94 ºC 20 seconds
61ºC 60 seconds
Drop -0,6 ºC per cycle
] 10 cycles

94 ºC 10 seconds
55ºC 60 seconds
] 25 cycles

20ºC ∞