A predictive model for canine dilated cardiomyopathy – a meta-analysis of Doberman Pinscher data

Siobhan Simspon, Jennifer Edwards, Richard D Emes, Malcolm A Cobb, Nigel P Mongan, Catrin S Rutland

Dilated cardiomyopathy is a prevalent and often fatal disease in humans and dogs. Indeed dilated cardiomyopathy is the third most common form of cardiac disease in humans, reported to affect approximately 36 individuals per 100,000 individuals. In dogs, dilated cardiomyopathy is the second most common cardiac disease and is most prevalent in the Irish Wolfhound, Doberman Pinscher and Newfoundland breeds. Dilated cardiomyopathy is characterised by ventricular chamber enlargement and systolic dysfunction which often leads to congestive heart failure. Although multiple human loci have been implicated in the pathogenesis of dilated cardiomyopathy, the identified variants are typically associated with rare monogenic forms of dilated cardiomyopathy. The potential for multigenic interactions contributing to human dilated cardiomyopathy remains poorly understood. Consistent with this, several known human dilated cardiomyopathy loci have been excluded as common causes of canine dilated cardiomyopathy, although canine dilated cardiomyopathy resembles the human disease functionally. This suggests additional genetic factors contribute to the dilated cardiomyopathy phenotype.

This study represents a meta-analysis of available canine dilated cardiomyopathy genetic datasets with the goal of determining potential multigenic interactions relating the sex chromosome genotype (XX vs XY) with known dilated cardiomyopathy associated loci on chromosome 5 and the PDK4 gene in the incidence and progression of dilated cardiomyopathy. The results show an interaction between known canine dilated cardiomyopathy loci and an unknown X-linked locus. Our study is the first to test a multigenic contribution to dilated cardiomyopathy and suggest a genetic basis for the known sex-disparity in dilated cardiomyopathy outcomes.

Original Article

A predictive model for canine dilated cardiomyopathy – a meta-analysis of Doberman Pinscher data

Siobhan Simpson[a], Jennifer Edwards[a], Richard D. Emes[a,b], Malcolm Cobb[a*], Nigel P. Mongan[a,c*] and Catrin S. Rutland[a].

[a] *School of Veterinary Medicine and Science, Sutton Bonington Campus, University of Nottingham, College Road, Leicestershire, LE12 5RD, UK.*

[b] *Advanced Data Analysis Centre, Sutton Bonington Campus, University of Nottingham, College Road, Leicestershire, LE12 5RD, UK.*

[c] *Department of Pharmacology, Weill Cornell Medical College, 1300 York Ave., New York, NY, 10065, USA*

* Corresponding author. Tel.: 0044 (0) 115 95 16625 *School of Veterinary Medicine and Science, Sutton Bonington Campus, University of Nottingham, College Road, Leicestershire, LE12 5RD, UK.*

E-mail address: Nigel.mongan@nottingham.ac.uk / npm2001@med.cornell.edu

Introduction

Dilated cardiomyopathy (DCM) is a prevalent and often fatal disease requiring clinical management in humans and dogs (Egenvall, Bonnett & Häggström, 2006; Hershberger, Morales & Siegfried, 2010). DCM is the second most common cardiac disease in dogs and is characterised by ventricular chamber enlargement and systolic dysfunction which often leads to congestive heart failure(Egenvall, Bonnett & Häggström, 2006). The aetiology of DCM is complex. Genetic factors, myocardial ischemia, hypertension, toxins, infections and metabolic defects have been implicated (McNally, Golbus & Puckelwartz, 2013). To date mutations in over 50 genes have been associated with DCM in humans, however mutations in the most prevalent DCM related genes only account for approximately 50% of patients with DCM (Posafalvi et al., 2012). In human DCM genetic testing where a panel of approximately 50 loci are tested concurrently, often more than one locus can be implicated in the disease (McNally, Golbus & Puckelwartz, 2013), suggesting multiple genetic factors cooperate in DCM aetiology.

Canine DCM is phenotypically similar to human DCM (Shinbane et al., 1997). As outlined below, to date mutations in only two genes (*PDK4* and *STRN*) and a single nucleotide polymorphism (SNP) on chromosome 5 have been associated with canine DCM (Mausberg et al., 2011; Meurs et al., 2012, 2013), suggesting additional genetic causes remain unknown. While canine studies have sometimes been limited by small sample size (typically less than 10 individuals), those studies with larger sample numbers (greater than 50 individuals) have also frequently failed to find significant associations with DCM (e.g. Philipp et al. 2007, 2008; Wiersma et al. 2008). One possible explanation for the challenges in identifying DCM associated loci in humans and dogs is that even within an extended family or breed, genetic variation at a single locus cannot explain the development of DCM. Indeed dog breeds can be considered as large families, with dogs within a breed more related to each other than dogs of other breeds (Parker et al., 2004). In the same way that some human families are

affected by DCM, a subset of dog breeds are affected by DCM more frequently than others (Egenvall, Bonnett & Häggström, 2006). Dobermans Pinschers (hereafter Dobermans) are particularly affected by DCM, with both a high prevalence (58.2% in European Dobermans) and severity with DCM associated death often occurring within 8 weeks of diagnosis (Calvert et al., 1997; Wess et al., 2010). In dogs, diagnosis is usually at the onset of clinical symptoms of heart failure. But there is an extended pre-clinical phase, during which if treatment can be effective by prolonging the onset of heart failure (Summerfield et al., 2012). In this phase left ventricular dilation and dysfunction begins, and can be accompanied by ventricular premature complexes (Singletary et al., 2012), Median life expectancy of DCM affected European Dobermans is 7.8 years, compared with 11 years for unaffected European Dobermans (Proschowsky, Rugbjerg & Ersbøll, 2003; Egenvall, Bonnett & Häggström, 2006). A deletion in a splice site of the *PDK4* gene (Meurs et al., 2012) and a SNP on chromosome 5 (Mausberg et al., 2011) in Dobermans are two of only three canine DCM mutations identified to date. While these two loci are associated with Doberman DCM, individually neither locus explains all cases of Doberman DCM (Mausberg et al., 2011; Meurs et al., 2012). Individuals heterozygous at the Chr5 SNP are more likely to develop DCM, but there are many DCM cases homozygous for the healthy allele (Mausberg et al., 2011). While *PDK4* genotypes are less definite predictors of DCM, with both affected and unaffected individuals possessing the three possible genotypes, the 16bp *PDK4* splice site deletion is found more frequently in North American Dobermans with DCM than those without DCM (Meurs et al., 2012). However an analysis of European Dobermans failed to identify an association between *PDK4* and DCM (Owczarek-Lipska et al., 2013), suggesting additional unknown factors influence the effect of *PDK4* in predisposing individuals to DCM. Thus novel genetic causes of canine DCM remain to be identified (Mausberg et al., 2011; Philipp et al., 2012).

In this study we developed genetic models to test the influence of unknown genetic factors to predict which DCM-associated genotype combinations are likely to develop DCM. Using this method

76 we provide evidence for a sex-linked genetic influence on known DCM loci in the pathogenesis of
77 canine DCM. Our study is the first to propose a multigenic contribution to canine DCM.
78

79 **Materials and Methods**

80 *Model development*

81 A literature search of the Pubmed and Web of Science database using the following search
82 terms: “Doberman DCM loci”, “Doberman Dilated Cardiomyopathy loci”, “Doberman DCM gene”,
83 “Doberman Dilated Cardiomyopathy gene”, “Doberman DCM locus”, “Doberman Dilated
84 Cardiomyopathy locus” (Figure 1) identified thirty unique records. This search identified two loci
85 associated with DCM in North American Dobermans (Meurs et al., 2012**;** Mausberg et al., 2011). By
86 combining the genotypes from the identified Doberman DCM associated loci, and additional putative
87 loci, predictive models were developed and tested against observed DCM incidence data. All genotype
88 combinations for the DCM associated SNP identified on chromosome 5
89 (TIGRP2P73097:CFA5:g.53,941,386T>C, CanFam2.1) (Mausberg et al., 2011) and the *PDK4*
90 (GeneID:482310) splice site deletion (CFA14:g.20,829,667_20,829,682del, CanFam3.1) (Meurs et al.,
91 2012) were determined. Further analysis determined which genotype combinations were likely to lead
92 to DCM. Some genotypes are definitive; all individuals homozygous for the susceptibility allele at
93 CFA5:g.53,941,386T>C develop DCM (Mausberg et al., 2011).
94

95 *Determining which genotypes develop DCM*

96 Five genetic models incorporating genotypes at multiple observed and hypothetical loci were
97 developed including: 1. two known DCM loci; 2. two known loci + 50% of the population more
98 susceptible to developing DCM; 3. two known loci + a novel autosomal dominant DCM locus; 4. two

known loci + a novel autosomal recessive DCM locus; 5. two known loci + a novel additive DCM locus and 6. two known loci + a novel X-linked DCM locus. For each model, different biologically feasible phenotype outcomes were tested for each genotype combination to establish the best fit of the model to the observed DCM incidence data. Each model was subject to the following constraints: individuals that are homozygous CC at the Chr5 SNP develop DCM, and individuals with no susceptibility alleles are healthy.

Model testing

For each model the frequency of each genotype combination was calculated by multiplying the genotype frequencies using *PDK4* and Chr5 frequencies (Table 1) obtained from Owczarek-Lipska et al. (2013) and Mausberg et al. (2011). A range of frequencies were tested for each hypothetical loci. For example, for the model incorporating only *PDK4* and Chr5 variants, one genotype combination is WtWt-TT. The frequency of this genotype combination is the product of the frequency of WtWt and the frequency of TT in the population. From the combined genotype frequencies the expected numbers of individuals with each genotype combination were calculated by multiplying the frequency by the number of individuals in the study to be compared with (182 when compared with Mausberg, et al. (2011) and Owczarek-Lipska et al. (2013)). Thus the numbers of individuals in the model that were, for example, WtWt healthy and WtWt DCM were obtained by summing the numbers in each category. Having obtained the numbers of affected and unaffected individuals that the model predicts for each genotype, these were tested against the observed data using a χ^2 test. Where additional putative DCM loci were included in the model, several allele frequencies were tested. However, as GWAS studies have previously been carried out (Mausberg et al., 2011; Meurs et al., 2012) it is unlikely that additional DCM alleles are at higher frequencies than those already identified. For this reason DCM

allele frequencies over 0.5 were not tested. If the model is a good fit of the observed data the χ^2 test statistic will be non-significant.

The proportion of the population that the model predicts to have DCM was determined by taking the sum of all the genotype combined frequencies that lead to DCM in the model. For example, for the model incorporating just the two known loci this is 0.0144+0.0624+0.0052+0.0048+0.0004 = 0.0872 – (Supplementary Table 1). This proportion was then compared to the observed DCM frequency of 0.582 (Wess et al., 2010).

For most models it must be assumed that there is no difference in DCM incidence between the sexes, as an effect of sex has not been included. For the DCM model testing a 50% increased susceptibility, where it is biologically feasible that males are more susceptible and the models incorporate an additional X-linked locus, it is possible to calculate the proportion males and females that develop DCM. While males develop clinical symptoms earlier and appear to be more severely affected, there are indications that the sex of those affected by DCM is close to 50% male, 50% female (Wess et al., 2010), so we would expect our model to reflect this.

Odds ratios of each genotype and allele developing DCM for each model were obtained by testing each genotype against the other two combined and each allele against the other. Odds ratios are the odds/probability of an individual with a particular genotype or allele developing DCM compared, by dividing one by the other, to the odds of an individual with all other genotypes or alleles developing DCM, with and odds ratio greater than one associated with the trait of interest and an odds ratio of less than one not associated (Bland & Altman, 2000). For example the odds ratio for TT in the published data from Mausberg et al. (2011) is calculated in the following way. There are 45 individuals that are TT DCM and 85 TT healthy the odds of a TT individual developing DCM are 45/85 (0.53), there are 43 individuals which are TC or CC with DCM and 9 individuals that are TC or CC healthy so the odds

of these individuals developing DCM are 43/9 (4.78) the odds ratio divides the genotype of interest odds by the 'others' odds to give the odds ratio or 0.11. To assess the significance of these ratios χ^2 tests were performed on the 2x2 tables – in the above example the four groups are TT-DCM, TT-healthy, TC or CC-DCM, TC or CC-healthy. If the model is a good fit to the observed data it is expected that the odds ratios are of a similar pattern and significance, e.g. TT, small – significantly not associated with DCM; TC, large – significantly associated with DCM; CC, not possible to test – not testable, as for the Chr5 SNP in Table 2. Odds ratios of both genotypes and alleles were obtained from the original studies (Tables 2 and 3).

Results

Following the constraints stated in the methods and using biologically feasible reasoning each model was optimised to best fit the observed data. For each model the genotype-phenotype decision descriptions are shown in Table 4. Tables of each model are in supplementary material.

Comparing model predictions with observed data

χ^2 test values comparing predicted numbers with observed numbers of DCM and healthy individuals at each genotype ranged from 4.35 to 7766.06. A χ^2 value of less than 11.07 indicates there is no significant difference between predicted and observed genotype-phenotype data, (5% significance level, with 5 degrees of freedom). Values less than 15.09 represent predictions not significantly different to observed values at the 1% significance level. χ^2 values less than these critical values are indicated in Table 5.

168 *Model predicted DCM population frequency and sex incidence*

169 For each model the predicted DCM frequency was calculated to provide an additional method
170 to test the accuracy of the model. The DCM frequency in the European Doberman population is
171 estimated to be 58.2% (Wess et al., 2010) , therefore accurate models should predict a similar
172 frequency. The frequencies predicted by each model are displayed in Table 6 (see also Supplementary
173 Table2), with those within 10% of the reported frequency highlighted as accurate models. Further to
174 this the proportion of males and females that each model predicts to develop DCM were calculated.
175 Whilst most models do not account for sex and assume equal numbers of males and females affected,
176 two models tested either a 50% increase in male susceptibility or an additional X-linked locus. Based
177 on reported DCM incidence for a model to fit the observed data it is expected that similar proportions
178 of males and females develop DCM. Table 7 shows that irrespective of the frequency of the novel
179 susceptibility allele the model incorporating a novel X linked DCM locus gives similar proportions of
180 affected males and females.

181

182 *Odds ratios*

183 For the Chr5 SNP there are no odds ratio for CC as all individuals that are CC develop DCM in
184 both the original study (Mausberg et al., 2011) and models so odds ratios cannot be calculated. Despite
185 this a χ^2 test can be performed on the counts of affected and unaffected individuals observed and
186 predicted with the genotype so the significance of the results was obtained. For the Chr5 SNP, 12 of 18
187 models (Table 9), and 15 of the allele odds ratios are consistent with the original studies (Table 11).
188 The *PDK4* deletion association was identified in the North American Doberman population, in the
189 European population the odds ratios (Tables 8 & 10) are not significantly different from the null result
190 of 1. Once combined with additional loci similar significant likelihood ratios as the North American
191 population are obtained for 13 of 18 models (Tables 8 & 10).

192

193 *Selecting the most realistic model*

194 For a model to be considered plausible it should predict similar numbers of affected and
195 unaffected individuals at each genotype as observed in Mausberg et al. (2011) and Owczarek-Lipska et
196 al. (2013), predict similar DCM frequency as reported in the population (Wess et al., 2010), and give
197 odds ratios of genotypes and alleles similar to those from the studies which report an association. To
198 assist in determining which models meet these requirements Table 12 shows which conditions each
199 model meets (Supplementary Table 3-6). From this it is possible to see that no model meets all the
200 conditions, but two similar models, the models incorporating the two identified loci and an additional
201 X-linked DCM locus with the novel DCM allele frequency at 0.4 and 0.5, meet all but one condition
202 each. An additional exploration of the additional X-linked DCM allele frequency indicates that an X-
203 linked DCM allele frequency between 0.4 and 0.5 leads to all conditions being met.

204

Discussion

This study used publicly available data to test the prediction that genetic models incorporating multiple factors can better explain and predict the incidence of canine DCM than those utilising a single factor. Until now, the possibility that multiple genes combine to influence DCM phenotype has been proposed, but has not yet been tested, despite an established role for multiple loci in related diseases (Ingles et al., 2005; Xu et al., 2010; Rampersaud et al., 2011; Posafalvi et al., 2012). This is the first study to investigate the combined effect of multiple factors on the predisposition to DCM. Although our models do not explain all cases of canine DCM, by combining three factors (*PDK4*, Chr5 TIGRP2P73097 SNP and an X-linked locus) we show that DCM incidence can be more accurately predicted (Tables 6-12). Furthermore, as noted above the *PDK4* splice site deletion is not significantly associated with DCM in the European population. But in the model incorporating only the two known loci, the *PDK4* variant improves the odds ratio for the Chr5 SNP. Collectively these findings indicate that models incorporating multiple factors are more effective than those incorporating a single factor. This result is important because it has implications for future studies of the genetics and management of DCM. A better understanding of the genetic basis of DCM will permit the monitoring and earlier clinical intervention of high risk individuals thus potentially improving the outcome for affected individuals.

To assess the accuracy of each model we performed several statistical tests. For any model to be considered an accurate representation of observed data it should predict similar numbers of affected and unaffected individuals at each genotype as have been reported in the published data. It should also predict a similar DCM frequency to that found in the population. Secondly, the odds ratios of genotypes and alleles should support an association of the specific variants with DCM. The models incorporating the two known DCM loci and an additional X-linked locus with a susceptible allele frequency of 0.46 for the novel susceptible allele met all such conditions. It is important to note that

this susceptible allele frequency should have been identified by the previous GWAS studies (Mausberg et al., 2011; Meurs et al., 2012). It is therefore possible that additional cases and controls are required to complete a comprehensive GWAS analysis of DCM in Dobermans to establish the function and frequency of this predicted DCM associated locus.

Most predictive models are based on either known or simulated genotypes at multiple loci (Janssens et al., 2006; Pencina, D'Agostino & Vasan, 2008). Such models do not account for known effects of genotypes or allow the inclusion of additional as yet unknown, loci. For example, in this study all individuals possessing the Chr5 CC genotype have DCM. Our methodology is unique and useful where there are multiple known and unknown factors which do not fully account for the phenotype. In particular our approach accommodates specific gene combinations to lead to disease, rather than incremental risk factors as is the case in other predictive models (Janssens et al., 2006; Pencina, D'Agostino & Vasan, 2008). Limitations to our methodology include the number of factors that can be modelled is limited by the available data. Despite this our methodology could be used in other situations. While many phenotypes are the consequence of multiple loci, there can be some loci which have comparatively more important contribution to the phenotype (e.g. Strange et al. 2011; Papa et al. 2013). Identifying these loci can be the first steps in predicting phenotypes (e.g. Hayes et al. 2010; Papa et al. 2013). Following the identification of loci associated with a trait, our methodology can be used to indicate what type of additional loci may be influencing the trait of interest, which may simplify the identification of additional loci.

Conclusions

There are many unknown factors involved in the aetiology of canine and human DCM. In Dobermans we have identified multigenic effects and a possible X-linked locus as novel variables influencing

DCM risk. While the *PDK4* splice site deletion and the Chr5 SNP have both been tested for association with DCM in the European population of Dobermans, the combined genotype of individuals has not yet been considered (Mausberg et al., 2011; Owczarek-Lipska et al., 2013). Our model would benefit from further genotyping of Dobermans at both the *PDK4* and Chr5 variants to further validate the model. Future work is also required to identify X-linked DCM loci if the model is verified for the known loci. It is also possible that the different combinations of alleles leading to DCM in the model could affect the time taken to progress from one disease stage to the next as reported by Wess et al. (2010). If validated our model has implications for current canine breeding practices and welfare of individuals within the breed. Individuals with allele combinations more likely to develop DCM can be monitored more intensely than those with less genetic risk and mating pairs resulting in deleterious genotypes can be avoided. This will have improve welfare by reducing the prevalence of DCM-associated alleles within the population and potentially improving the longevity of affected dogs by enabling monitoring and earlier clinical management. By utilising similar methodology equivalent multigenic effects and possible additional loci could be identified in human DCM, giving similar benefits to those described for Dobermans.

Conflict of interest statement

None of the authors of this paper has a financial or personal relationship with other people or organisations that could inappropriately influence or bias the content of the paper.

References:

Bland JM, Altman DG. 2000. Statistics notes. The odds ratio. *BMJ (Clinical research ed.)* 320:1468.

Calvert CA, Pickus CW, Jacobs GJ, Brown J. 1997. Signalment, Survival, and Prognostic Factors in Doberman Pinschers With End-Stage Cardiomyopathy. *Journal of Veterinary Internal Medicine* 11:323–326.

Egenvall A, Bonnett BN, Häggström J. 2006. Heart Disease as a Cause of Death in Insured Swedish Dogs Younger Than 10 Years of Age. *Journal of Veterinary Internal Medicine* 20:894–903.

Hayes BJ, Pryce J, Chamberlain AJ, Bowman PJ, Goddard ME. 2010. Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLoS genetics* 6:e1001139.

Hershberger RE, Morales A, Siegfried JD. 2010. Clinical and genetic issues in dilated cardiomyopathy: a review for genetics professionals. *Genetics in medicine : official journal of the American College of Medical Genetics* 12:655–67.

Ingles J, Doolan A, Chiu C, Seidman J, Seidman C, Semsarian C. 2005. Compound and double mutations in patients with hypertrophic cardiomyopathy: implications for genetic testing and counselling. *Journal of medical genetics* 42:e59.

Janssens ACJW, Aulchenko YS, Elefante S, Borsboom GJJM, Steyerberg EW, van Duijn CM. 2006. Predictive testing for complex diseases using multiple genes: Fact or fiction? *Genetics in Medicine* 8:395–400.

Mausberg T-B, Wess G, Simak J, Keller L, Drögemüller M, Drögemüller C, Webster MT, Stephenson H, Dukes-McEwan J, Leeb T. 2011. A Locus on Chromosome 5 Is Associated with Dilated Cardiomyopathy in Doberman Pinschers. *PLoS ONE* 6:e20042.

McNally EM, Golbus JR, Puckelwartz MJ. 2013. Genetic mutations and mechanisms in dilated cardiomyopathy. *The Journal of Clinical Investigation* 123:19–26.

Meurs KM, Lahmers S, Keene BW, White SN, Oyama MA, Mauceli E, Lindblad-Toh K. 2012. A splice site mutation in a gene encoding for PDK4, a mitochondrial protein, is associated with the development of dilated cardiomyopathy in the Doberman pinscher. *Human Genetics* 131:1319–1325.

Meurs KM, Stern JA, Sisson DD, Kittleson MD, Cunningham SM, Ames MK, Atkins CE, DeFrancesco T, Hodge TE, Keene BW et al. 2013. Association of dilated cardiomyopathy with the striatin mutation genotype in boxer dogs. *Journal of veterinary internal medicine / American College of Veterinary Internal Medicine* 27:1437–40.

Owczarek-Lipska M, Mausberg T-B, Stephenson H, Dukes-McEwan J, Wess G, Leeb T. 2013. A 16-bp deletion in the canine PDK4 gene is not associated with dilated cardiomyopathy in a European cohort of Doberman Pinschers. *Animal genetics* 44:239.

Papa R, Kapan DD, Counterman BA, Maldonado K, Lindstrom DP, Reed RD, Nijhout HF, Hrbek T, McMillan WO. 2013. Multi-allelic major effect genes interact with minor effect QTLs to control adaptive color pattern variation in Heliconius erato. *PloS one* 8:e57033.

Parker HG, Kim L V, Sutter NB, Carlson S, Lorentzen TD, Malek TB, Johnson GS, DeFrance HB, Ostrander EA, Kruglyak L. 2004. Genetic structure of the purebred domestic dog. *Science* 304:1160–1164.

Pencina MJ, D'Agostino RB, Vasan RS. 2008. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine* 27:157–72; discussion 207–12.

Philipp U, Broschk C, Vollmar A, Distl O. 2007. Evaluation of tafazzin as candidate for dilated cardiomyopathy in Irish wolfhounds. *J Hered* 98:506–509.

Philipp U, Vollmar A, Distl O. 2008. Evaluation of six candidate genes for dilated cardiomyopathy in Irish wolfhounds. *Anim Genet* 39:88–89.

Philipp U, Vollmar A, Häggström J, Thomas A, Distl O. 2012. Multiple Loci Are Associated with Dilated Cardiomyopathy in Irish Wolfhounds. *PLoS ONE* 7:e36691.

Posafalvi A, Herkert JC, Sinke RJ, van den Berg MP, Mogensen J, Jongbloed JDH, van Tintelen JP. 2012. Clinical utility gene card for: dilated cardiomyopathy (CMD). *European Journal of Human Genetics*.

Proschowsky HF, Rugbjerg H, Ersbøll AK. 2003. Mortality of purebred and mixed-breed dogs in Denmark. *Preventive Veterinary Medicine* 58:63–74.

Rampersaud E, Siegfried JD, Norton N, Li D, Martin E, Hershberger RE. 2011. Rare variant mutations identified in pediatric patients with dilated cardiomyopathy. *Progress in pediatric cardiology* 31:39–47.

Shinbane JS, Wood MA, Jensen DN, Ellenbogen KA, Fitzpatrick AP, Scheinman MM. 1997. Tachycardia-Induced Cardiomyopathy: A Review of Animal Models and Clinical Studies. *Journal of the American College of Cardiology* 29:709–715.

Singletary GE, Morris NA, Lynne O'Sullivan M, Gordon SG, Oyama MA. 2012. Prospective evaluation of NT-proBNP assay to detect occult dilated cardiomyopathy and predict survival in Doberman Pinschers. *Journal of veterinary internal medicine / American College of Veterinary Internal Medicine* 26:1330–6.

Strange A, Li P, Lister C, Anderson J, Warthmann N, Shindo C, Irwin J, Nordborg M, Dean C. 2011. Major-effect alleles at relatively few loci underlie distinct vernalization and flowering variation in Arabidopsis accessions. *PloS one* 6:e19949.

Summerfield NJ, Boswood A, O'Grady MR, Gordon SG, Dukes-McEwan J, Oyama MA, Smith S, Patteson M, French AT, Culshaw GJ et al. 2012. Efficacy of pimobendan in the prevention of congestive heart failure or sudden death in Doberman Pinschers with preclinical dilated cardiomyopathy (the PROTECT Study). *Journal of veterinary internal medicine / American College of Veterinary Internal Medicine* 26:1337–49.

Wess G, Schulze A, Butz V, Simak J, Killich M, Keller LJM, Maeurer J, Hartmann K. 2010. Prevalence of dilated cardiomyopathy in Doberman Pinschers in various age groups. *Journal of Veterinary Internal Medicine* 24:533–538.

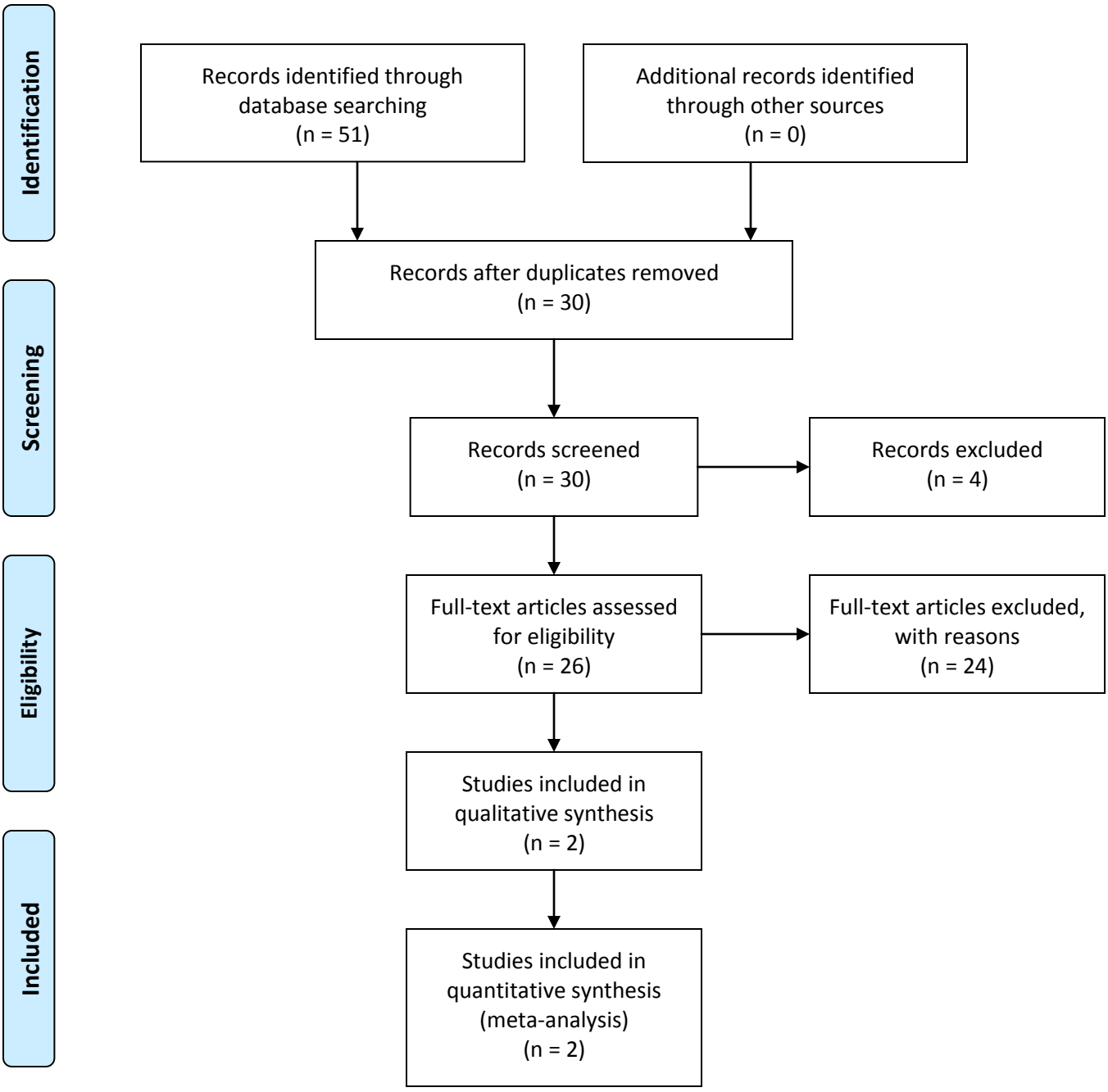
Wiersma AC, Stabej P, Leegwater PAJ, Van Oost BA, Ollier WE, Dukes-McEwan J. 2008. Evaluation of 15 Candidate Genes for Dilated Cardiomyopathy in the Newfoundland Dog. *Journal of Heredity* 99:73–80.

Xu T, Yang Z, Vatta M, Rampazzo A, Beffagna G, Pilichou K, Pillichou K, Scherer SE, Saffitz J, Kravitz J et al. 2010. Compound and digenic heterozygosity contributes to arrhythmogenic right ventricular cardiomyopathy. *Journal of the American College of Cardiology* 55:587–97.

Figure 1(on next page)

Prisma Flow Diagram

PRISMA 2009 Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Table 1(on next page)

Genotype frequencies assuming Hardy Weinberg Equilibrium

Allele frequencies taken from Mausberg et al. (2011) and Owczarek-Lipska et al. (2013)

Table 1. Genotype frequencies assuming Hardy Weinberg Equilibrium

PDK4		***Chr5* SNP**	
genotype	**freq**	**genotype**	**freq**
Wt Wt	0.72	TT	0.74
Wt del	0.26	TC	0.24
Del del	0.02	CC	0.02

Allele frequencies taken from Mausberg et al. (2011) and Owczarek-Lipska et al. (2013)

Table 2 (on next page)

Genotype odds ratios from the original studies reporting an association

Ratios from the PDK4 locus (Meurs et al., 2012) and Chromosome 5 SNP (Mausberg et al., 2011). The PDK4 χ2 test results indicate that the WtWt genotype significantly associated with non-DCM and the WtDel genotype significantly associated with DCM at the 0.01 significance level, the DelDel genotype odds ratio whilst different from the null result of 1, is not significantly so. For the chromosome 5 SNP all individuals that are CC in the original study developed DCM, thus and odds ratio and confidence interval cannot be calculated, but χ2 tests can be performed on the data. TT is significantly associated with non-DCM and the TC and CC genotypes are significantly associated with DCM at the 0.01 significance level.

Table 2. Genotype odds ratios from the original studies reporting an association.

Genotype	Odds ratio	95% CI
PDK4 WtWt	0.14	0.07, 0.32
PDK4 WtDel	5.21	2.70, 12.09
PDK4 DelDel	1.14	0.41, 3.18
Chr5 TT	0.11	0.05, 0.24
Chr5 TC	6.23	2.78, 14.00
Chr5 CC	NA	NA

Ratios from the PDK4 locus (Meurs et al., 2012) and Chromosome 5 SNP (Mausberg et al., 2011). The PDK4 χ^2 test results indicate that the WtWt genotype significantly associated with non-DCM and the WtDel genotype significantly associated with DCM at the 0.01 significance level, the DelDel genotype odds ratio whilst different from the null result of 1, is not significantly so. For the chromosome 5 SNP all individuals that are CC in the original study developed DCM, thus and odds ratio and confidence interval cannot be calculated, but χ^2 tests can be performed on the data. TT is significantly associated with non-DCM and the TC and CC genotypes are significantly associated with DCM at the 0.01 significance level.

Table 3(on next page)

Allele odds ratios from the original studies reporting DCM associations

Allele odds rations at the PDK4 locus (Meurs et al., 2012) and Chromosome 5 SNP (Mausberg et al., 2011). The χ2 test results indicate that each susceptibility (Del and C respectively) allele is significantly associated with DCM and the alternate allele significantly associated with non-DCM at the 0.01 significance level.

Table 3. Allele odds ratios from the original studies reporting DCM associations.

Allele	Odds ratio	95% CI
PDK4 Wt	0.38	0.23, 0.64
PDK4 Del	2.63	1.57, 4.42
Chr5 T	0.12	0.06, 0.26
Chr5 C	8.11	3.85, 17.09

Allele odds rations at the PDK4 locus (Meurs et al., 2012) and Chromosome 5 SNP (Mausberg et al., 2011). The χ^2 test results indicate that each susceptibility (Del and C respectively) allele is significantly associated with DCM and the alternate allele significantly associated with non-DCM at the 0.01 significance level.

Table 4(on next page)

Genotype-phenotype decision descriptions for each model

Models represent: 1. the two known DCM loci; 2. two known loci + 50% of the population is more susceptible to developing DCM; 3. two known DCM loci combined with a novel autosomal dominant DCM locus; 4. two known DCM loci combined with an autosomal recessive locus; 5. two known DCM loci combined with a an additional DCM locus that is additive and 6. two known DCM loci combined with an X-linked DCM locus

Table 4. Genotype-phenotype decision descriptions for each model.

Model	Genotype – phenotype decision description, in addition to the rules:
1	DCM develops when both the PDK4 locus and Chr5 SNP have at least one DCM susceptibility allele.
2	50% more susceptible only need to have a single DCM susceptibility allele at either locus to develop DCM while the 50% less susceptible to DCM require at least one DCM susceptibility allele at both loci to develop DCM.
3	All individuals that have a susceptibility allele at the additional locus develop DCM. Those individuals with no susceptibility alleles at the additional locus need at least one DCM susceptibility allele at both of the other loci to develop DCM.
4	All homozygous susceptible individuals at the additional locus develop DCM. For individuals that are heterozygous at the additional locus, DCM occurs when combined with another DCM susceptibility allele, while homozygous unsusceptible individuals need at least one DCM susceptibility allele at both of the other loci to develop DCM.
5	All homozygous susceptible individuals at the additional locus develop DCM. Heterozygotes and homozygous unsusceptible individuals need at least one DCM susceptibility allele at both of the other loci to develop DCM.
6	X linked susceptible DCM locus males can either possess a single unsusceptible X (XY) or a single susceptible x (xY), while females can be unsusceptible X homozygotes (XX), heterozygotes (Xx) or susceptible x homozygotes (xx). Unsusceptible X males (XY) are phenotypically identical to unsusceptible X homozygotes (XX) with these

	individuals requiring at least one DCM susceptibility allele at both of the other loci to develop DCM. All individuals that possess a susceptible X (xY and xx individuals) develop DCM in this model while heterozygotes (Xx) only require a single DCM susceptibility allele at one of the other loci to develop DCM.

Models represent: 1. the two known DCM loci; 2. two known loci + 50% of the population is more susceptible to developing DCM; 3. two known DCM loci combined with a novel autosomal dominant DCM locus; 4. two known DCM loci combined with an autosomal recessive locus; 5. two known DCM loci combined with a an additional DCM locus that is additive and 6. two known DCM loci combined with an X-linked DCM locus

Table 5(on next page)

χ2 test statistic results comparing predicted of DCM and healthy individuals at each genotype from each model with observed numbers of DCM and healthy individuals at each genotype.

Model data based on data from Mausberg et al. (2011) – Chr5 SNP and Owczarek-Lipska et al. (2013) – PDK4. ** not significant at 5% significance level, * not significant at 1% significance level.

Table 5. χ^2 test statistic results comparing predicted of DCM and healthy individuals at each genotype from each model with observed numbers of DCM and healthy individuals at each genotype.

	χ^2 test statistic for each model							
Model	***PDK4***				**Chr5**			
1.	1269.23				7766.06			
2.	110.45				596.68			
DCM allele freq	**0.5**	**0.4**	**0.3**	**0.2**	**0.5**	**0.4**	**0.3**	**0.2**
3.	32.47	29.25	51.42	113.35	6.58**	7.69**	24.30	69.27
4.	26.24	74.61	171.69	379.06	31.65	67.45	145.76	360.86
5.	88.95	31.36	4.97**	4.36**	114.72	53.10	23.13	17.21
DCM X allele (x) freq	**0.5**	**0.4**	**0.3**	**0.2**	**0.5**	**0.4**	**0.3**	**0.2**
6.	10.57**	10.06**	25.38	71.30	11.32*	9.29**	19.55	52.86

Model data based on data from Mausberg et al. (2011) – Chr5 SNP and Owczarek-Lipska et al. (2013) – *PDK4*. ** not significant at 5% significance level, * not significant at 1% significance level.

Table 6(on next page)

DCM frequency predicted by each model

* indicates frequencies within 0.1 of the reported frequency (0.582 (Wess et al., 2010)) in the European Doberman pincher population.

Table 6. DCM frequency predicted by each model.

Model	**DCM freq for each model**			
1.	0.0872			
2.	0.2772			
DCM allele freq	**0.5**	**0.4**	**0.3**	**0.2**
3.	0.5054*	0.415648	0.328952	0.245321
4.	0.3154	0.233248	0.169352	0.123712
5.	0.7718	0.671392*	0.552728*	0.415808
DCM X allele (x) freq	**0.5**	**0.4**	**0.3**	**0.2**
6.	0.5245*	0.433984	0.350432	0.257536

* indicates frequencies within 0.1 of the reported frequency (0.582 (Wess et al., 2010)) in the European Doberman pincher population.

Table 7 (on next page)

Proportion of males and females predicted to be affected by DCM by models 2 and 6

Table 7. Proportion of males and females predicted to be affected by DCM by models 2 and 6

Model	Proportion DCM	
	Male	**Female**
2.	0.4672	0.0872
6.		
DCM X allele (x) freq 0.5	0.5436	0.5054
DCM X allele (x) freq 0.4	0.45232	0.415648
DCM X allele (x) freq 0.3	0.36104	0.339824
DCM X allele (x) freq 0.2	0.26976	0.245312

Table 8(on next page)

Odds ratios of each PDK4 genotype with χ2 significnace

** significant at 1% level, * significant at 5% level

Table 8. Odds ratios of each PDK4 genotype with χ^2 significnace.

	PDK4 genotype odds ratio											
Model	**wtwt**	**wtdel**	**deldel**	**wtwt**	**wtdel**	**deldel**	**wtwt**	**wtdel**	**deldel**	**wtwt**	**wtdel**	**deldel**
individual loci	0.78	1.29	1.11									
1.	0.06**	12.91**	3.85									
2.	0.1**	9.41**	4.6									
DCM allele freq	**0.5**			**0.4**			**0.3**			**0.2**		
3.	0.14**	6.70**	4.42	0.15**	6.31**	3.98	0.15**	6.21**	3.69	0.14**	6.47**	3.53
4.	0.45*	2.17*	1.76	0.35**	2.73**	2.03	0.25**	3.77**	2.43	0.15**	5.82**	2.98
5.	0.7	1.42	1.31	0.67	1.49	1.36	0.62	1.6	1.43	0.53	1.84	1.58
DCM X allele (x) freq	**0.5**			**0.4**			**0.3**			**0.2**		
6.	0.31**	3.12**	2.4	0.30**	3.23**	2.41	0.28**	3.41**	2.45	0.24**	3.89**	2.59

** significant at 1% level, * significant at 5% level

Table 9(on next page)

Odds ratios of each Chr5 SNP genotype with χ2 significance

** significant at 1% level, * significant at 5% level

Table 9. Odds ratios of each Chr5 SNP genotype with χ^2 significance.

Model	**Chr5 genotype odds ratio**											
	TT	**TC**	**CC**	**TT**	**TC**	**CC**	**TT**	**TC**	**CC**	**TT**	**TC**	**CC**
individual loci	0.11**	6.23**	-**									
1.	0.02**	11.37**	-**									
2.	0.09**	9.23**	-**									
DCM allele freq	**0.5**			**0.4**			**0.3**			**0.2**		
3.	0.14**	6.74**	-	0.14**	6.34**	-*	0.13**	6.25**	-**	0.12**	6.56**	-**
4.	0.35**	2.33*	-**	0.25**	2.96**	-**	0.16**	4.13**	-**	0.08**	6.45**	-**
5.	0.67	1.51	-	0.61	1.57	-	0.54	1.7	-	0.44*	1.96	-*
DCM X allele (x) freq	**0.5**			**0.4**			**0.3**			**0.2**		
6.	0.29**	3.22**	-	0.27**	3.34**	-*	0.24**	3.55**	-**	0.19**	4.08**	-**

** significant at 1% level, * significant at 5% level

Table 10(on next page)

Odds ratios of each PDK4 allele with χ2 significance

** significant at 1% level, * significant at 5% level

Table 10. Odds ratios of each PDK4 allele with χ^2 significance.

Model	**PDK4 allele odds ratio**							
	Wt	**Del**	**Wt**	**Del**	**Wt**	**Del**	**Wt**	**Del**
individual loci	0.81	1.23						
1.	0.17**	5.84**						
2.	0.16**	6.22**						
DCM allele freq	**0.5**		**0.4**		**0.3**		**0.2**	
3.	0.19**	5.37**	0.2**	4.91**	0.22**	4.65**	0.22**	4.57**
4.	0.52*	1.94*	0.43**	2.32**	0.34**	2.94**	0.22**	3.91**
5.	0.74	1.36	0.71	1.36	0.66	1.51	0.59	1.69
DCM X allele (x) freq	**0.5**		**0.4**		**0.3**		**0.2**	
6.	0.37**	2.71**	0.36**	2.76**	0.35**	2.94**	0.32**	3.1**

** significant at 1% level, * significant at 5% level

Table 11(on next page)

Odds ratios of each Chr5 SNP allele with χ2 significance

** significant at 1% level, * significant at 5% level

Table 11. Odds ratios of each Chr5 SNP allele with χ^2 significnace.

Model	**Chr5 allele odds ratio**							
	T	**C**	**T**	**C**	**T**	**C**	**T**	**C**
individual loci	0.15**	6.64**						
1.	0.08**	12.33**						
2.	0.13**	7.49**						
DCM allele freq	**0.5**		**0.4**		**0.3**		**0.2**	
3.	0.19**	5.34**	0.19**	5.37**	0.18**	5.55**	0.16**	6.07**
4.	0.36**	2.76**	0.28**	3.62**	0.20**	5.08**	0.16**	7.68**
5.	0.72	1.38	0.64	1.38	0.55	1.82	0.45**	2.23**
DCM X allele (x) freq	**0.5**		**0.4**		**0.3**		**0.2**	
6.	0.33**	3.02**	0.3**	3.28**	0.27**	5.08**	0.23**	4.35**

** significant at 1% level, * significant at 5% level

Table 12(on next page)

Adherence to model

Table shows whether each model (with the new DCM allele frequency indicated) meets each condition, Y the condition is met, x the condition is not met. The number of conditions not met is also indicated.

2 **Table 12.** Adherence to model.

	χ^2			OR genotype		OR allele		
		Chr5			Chr5		Chr5	number of
model	**PDK4**	**SNP**	**DCM freq**	**PDK4**	**SNP**	**PDK4**	**SNP**	**conditions not met**
individual	-	-	-	x	Y	x	x	3
1.	x	x	x	Y	Y	Y	x	4
2.	x	x	x	Y	Y	Y	x	4
3.								
0.5	x	x	Y	Y	x	Y	Y	3
0.4	x	x	x	Y	Y	Y	Y	3
0.3	x	x	x	Y	Y	Y	Y	3
0.2	x	x	x	Y	Y	Y	Y	3
4.								
0.5	x	x	x	Y	Y	Y	Y	3
0.4	x	x	x	Y	Y	Y	Y	3
0.3	x	x	x	Y	Y	Y	Y	3
0.2	x	x	x	Y	Y	Y	Y	3
5.								
0.5	x	x	x	x	x	x	x	7
0.4	x	x	Y	x	x	x	x	6
0.3	Y	x	Y	x	x	x	x	5
0.2	Y	x	x	x	x	x	Y	5
6.								

0.5	Y	Y	Y	Y	x	Y	Y	1
0.4	Y	Y	x	Y	Y	Y	Y	1
0.3	x	x	x	Y	Y	Y	Y	3
0.2	x	x	x	Y	Y	Y	Y	3

Table shows whether each model (with the new DCM allele frequency indicated) meets each condition, Y the condition is met, x the condition is not met. The number of conditions not met is also indicated.