



24 **Abstract**

25 The diversity of eutherian reproductive strategies has led to variation in many traits, such  
26 as number of offspring, age of reproductive maturity, and gestation length. While  
27 reproductive trait variation has been extensively investigated and is well established in  
28 mammals, the genetic loci contributing to this variation remain largely unknown. The  
29 domestic dog, *Canis lupus familiaris* is a powerful model for studies of the genetics of  
30 inherited disease due to its unique history of domestication. To gain insight into the genetic  
31 basis of reproductive traits across domestic dog breeds, we collected phenotypic data for  
32 four traits – cesarean section rate (n = 97 breeds), litter size (n = 60), stillbirth rate (n =  
33 57), and gestation length (n = 23) – from primary literature and breeders’ handbooks. By  
34 matching our phenotypic data to genomic data from the Cornell Veterinary Biobank, we  
35 performed genome wide association analyses for these four reproductive traits, using body  
36 mass and kinship among breeds as co-variables. We identified 14 genome-wide significant  
37 associations between these traits and genetic loci, including variants near *CACNA2D3* with  
38 gestation length, *MSRB3* with litter size, *SMOC2* with cesarean section rate, *MITF* with litter  
39 size and still birth rate, *KRT71* with cesarean section rate, litter size, and stillbirth rate, and  
40 *HTR2C* with stillbirth rate. Some of these loci, such as *CACNA2D3* and *MSRB3*, have been  
41 previously implicated in human reproductive pathologies. Many of the variants that we  
42 identified have been previously associated with domestication-related traits, including  
43 brachycephaly (*SMOC2*), coat color (*MITF*), coat curl (*KRT71*), and tameness (*HTR2C*).  
44 These results raise the hypothesis that the artificial selection that gave rise to dog breeds  
45 also shaped the observed variation in their reproductive traits. Overall, our work

46 establishes the domestic dog as a system for studying the genetics of reproductive biology

47 and disease.

48

## 49 **Introduction**

50 Mammals exhibit wide variation in traits associated with reproduction (1-3). For example,  
51 gestation length can range from 12 days in the Gray dwarf hamster, *Cricetulus migratorius*,  
52 to 21 months in the African bush elephant, *Loxodonta africana* (4-6); neonate size can  
53 range from less than one gram in the shrew family (Soricidae), to more than a metric ton in  
54 the baleen whales (Balaenopteridae) (4,6); and neonates can be either precocial (e.g.,  
55 cricetid rodents, rabbits, and canids) or altricial (e.g., hystricomorph rodents, ungulates,  
56 and cetaceans) (1). This variation in reproductive traits also extends to methods of  
57 implantation (7), structure of the placenta (8,9), and lactation strategies (10,11). Not  
58 surprisingly, many reproductive traits also exhibit substantial intra-specific variation (5).  
59 For example, many mammals exhibit intraspecific variation in gestation length, including  
60 primates (12), rat and rabbits (13), as well as the domesticated cattle (14) and  
61 thoroughbred horses (15). Similarly, body fat percentages, which are associated with the  
62 energetics of reproduction, vary greatly between wild and captive baboons, and  
63 intraspecific variation among captive lemurs can vary from 8 - 41% (16).

64  
65 The existence of phenotypic variation in reproductive traits is well established, and can  
66 inform our understanding of the factors that shape patterns of survival and reproduction in  
67 both agricultural (17-20) and human populations (21). Not surprisingly, most genome  
68 wide association (GWAS) studies of reproductive traits focus on economically important  
69 traits in domesticated species, such as reproductive seasonality in rabbits (17), infertility in  
70 pigs (18), and dairy traits in cattle (19). GWAS studies focused on understanding human  
71 reproductive biology and its associated pathologies have also shed light on the genetic

72 basis of reproductive traits, including birth weight (22) and gestational duration or length  
73 (23-25). For example, maternal variation in six genomic loci (*ADCY5*, *AGTR2*, *EBF1*, *EEFSEC*,  
74 *RAP2C*, and *WNT4*) is associated with gestational duration and preterm birth (25). While  
75 these studies contribute to our understanding of the genetic architecture of reproductive  
76 traits, we still understand very little about the molecular pathways underlying this  
77 variation and are unable to explain the majority of the heritability in reproductive traits  
78 (26-29).

79  
80 To address this challenge, we studied the genetics of reproductive traits in a powerful new  
81 model system: the domestic dog. The dog is well-suited to this question, because the  
82 domestication bottleneck followed by intense artificial selection and inbreeding imposed  
83 over the past 300 years has led to the generation of more than 340 recognized breeds that  
84 exhibit dramatic morphological variation (30-32). Domestic dog breeds also show  
85 substantial variation in their reproductive traits. For example, Pomeranians and Norfolk  
86 Terriers typically have only 2 pups per litter, whereas Dalmatians and Rhodesian  
87 Ridgebacks typically sire 8-9 pups per litter (33). Similarly, 80 – 90% of French Bulldogs  
88 and Boston Terriers are born via cesarean section due to cephalopelvic disproportion,  
89 whereas only 2 – 3% of Australian Shepherds and Shar Peis require cesareans (34). Recent  
90 analyses have begun to study the genetic mechanisms that underlie the remarkable  
91 morphological variation between modern dog breeds in diverse traits such as snout length,  
92 ear erectness, and tail curliness (35-38), as well as genetic disease (39).

93

94 To gain insight into the genetic basis of reproductive traits across domestic dog breeds, we  
95 collected phenotypic data for four reproductive traits, namely cesarean section rate, litter  
96 size, stillbirth rate, and gestation length. We synthesized data from the primary literature  
97 and breeders' handbooks to obtain coverage of between 23 (trait) and 97 (trait) dog  
98 breeds, as well as body mass data from 101 dog breeds. By matching our phenotypic data  
99 to genome-wide genotypic data from the Cornell Veterinary Biobank, we performed GWAS  
100 analyses and identified 14 genetic loci that are significantly associated with these  
101 reproductive traits (using body mass as a co-variate). Several of these variants are in or  
102 near genes previously implicated in human reproduction-related pathologies. The majority  
103 of the variants that we discovered to be significantly associated with reproductive trait  
104 variation are also associated with domestication-related traits. For example, we found that:  
105 variation in a gene previously identified to be involved in brachycephaly is also  
106 significantly associated with rates of cesarean sections; variation in a gene previously  
107 associated with docility is also associated with stillbirth rates; and variation in genes  
108 previously linked to coat phenotypes, such as color and curliness, is also associated with  
109 several reproductive traits. These results suggest that selection for breed-specific  
110 morphological traits during dog domestication may have also directly or indirectly  
111 influenced variation in reproductive traits. More broadly, our results establish the domestic  
112 dog as a tractable system for studying the genetics of reproductive traits and underscore  
113 the potential for cryptic interactions between reproductive and other traits favored over  
114 the course of adaptation.

115

## 116 **Results**

117 To identify SNPs that are significantly associated with four reproductive traits in domestic  
118 dog breeds, we conducted across-breed GWAS analyses using a multivariate linear mixed  
119 model implemented in the program GEMMA (Zhou & Stephens, 2012). Number of  
120 individuals and distribution of breed varied with analysis. After filtering for MAF (MAF <  
121 0.05; 10,804 SNPs were excluded) and linkage disequilibrium (34,240 additional SNPs  
122 were excluded), 115,683 SNPs were included in the GWAS analysis for each reproductive  
123 trait. To validate our GWAS approach and analytical choices, we first used our collected  
124 values for body mass, a trait whose genetic associations have been previously extensively  
125 studied in dogs (36,37). As expected, our analysis recovered the major genes associated  
126 with dog breed body mass variation, including *IGF1* ( $P = 2.1 \times 10^{-31}$ ), *SMAD2*  
127 ( $P = 1.2 \times 10^{-17}$ ) and *IGF2BP2* ( $P = 5.1 \times 10^{-11}$ ) (Supplementary Figure 1,  
128 Supplementary Table 2).

129

### 130 **Four genetic loci significantly associate with cesarean section rate**

131 To examine whether there is variation in cesarean section rate among breeds, we first  
132 identified cesarean section rate values for a total of 97 of the 162 dog breeds with  
133 genotypic data (Supplementary Table 1). The cesarean section rate values were derived  
134 from a British survey across 151 breeds covering 13,141 bitches, which had whelped  
135 22,005 litters over the course of a 10 year period (34). The frequency of cesarean sections  
136 was estimated as the percentage of litters reported to be born by cesarean section. Among  
137 the 97 breeds with overlapping genetic data, the median cesarean section rate is 17.1%,  
138 with a minimum of 0% in Curly Coated Retrievers and Silky Terriers and a maximum of  
139 92.3% in Boston Terriers (Supplementary Figure 3A).

140

141 To identify SNPs that are significantly associated with the observed variation in cesarean  
142 section rate across domestic dog breeds, we conducted an across-breed GWAS analysis  
143 using 115,683 SNPs and cesarean section values across 95 dog breeds (Figure 1A,  
144 Supplementary Figure 2A). We identified four significant SNPs, three of which mapped to  
145 genes, namely paralemmin 3 (*PALM3*, uncorrected  $P = 1.4 \times 10^{-9}$ ), sparc-related  
146 modular calcium-binding protein 2 (*SMOC2*, uncorrected  $P = 2.0 \times 10^{-7}$ ), and keratin 71  
147 (*KRT71*, uncorrected  $P = 2.9 \times 10^{-7}$ ), and a fourth that mapped to the intergenic region  
148 between the *CD36* glycoprotein and a lincRNA (uncorrected  $P = 9.7 \times 10^{-8}$ ; Figure 1A).

149

150 The first significantly associated SNP (chromosome 1: 55,983,871) is found in the intron  
151 between exons 13 and 14 of *SMOC2*, a gene that is associated with brachycephaly in dogs  
152 (38,40); variation in *SMOC2* accounts for 36% of facial length variation in dogs (41). In  
153 humans, *SMOC2* is highly expressed in endometrium as well as other reproductive tissues,  
154 including the fallopian tubes, ovaries and cervix (Figure 2) (42). The 3' intronic location of  
155 the SNP raises the possibility that it might be regulatory (43).

156

157 The second SNP is found in the 3' UTR of *PALM3*, which is a member of the paralemmin  
158 gene family that also includes *PALM1*, *PALM2*, and *PALMD* (palmdelphin); members of this  
159 family are implicated in plasma membrane dynamics and as modulators of cellular cAMP  
160 signaling in the brain (44,45). The function of *PALM3* may be slightly different from the rest  
161 of the genes in the family, with recent work suggesting that *PALM3* is a binding protein of  
162 the single immunoglobulin IL-1 receptor-related molecule (SIGIRR), which is a negative



163 regulator of Toll-Interleukin-1 receptor signaling (46). In humans, *PALM3* is primarily  
164 expressed in the membranes of the stomach, kidney, parathyroid gland and epididymis  
165 (Figure 2) (42). The SNP (chromosome 20: 48,454,259) that is significantly associated with  
166 cesarean section rate is found in the first intron of the *PALM3* gene, suggesting that it might  
167 be involved in regulatory actions typically observed in 5' introns (43).

168

169 The third SNP (chromosome 27: 2,539,211) results in a missense mutation of exon 2 of  
170 *KRT71*, which belongs to a family of keratin genes specifically expressed in the inner root  
171 sheath of hair follicles (47). Prior analysis in dogs identified variation in gene *KRT71*, along  
172 with variation in genes *RSPO2* and *FGF5*, accounting for most coat phenotypes (48), such as  
173 curliness.

174

175 The fourth significant SNP (chromosome 18: 20,272,961) is found in the intergenic region  
176 between the *CD36* gene and a lincRNA (ENSCAFG00000034312). The protein product of  
177 CD36 is the fourth major glycoprotein of the platelet surface and serves as a receptor for  
178 thrombospondin in platelets (49). Other known functions include transport of long chain  
179 fatty acids (50).

180

### 181 **Six genetic loci significantly associate with litter Size**

182 To examine whether there are SNPs that are significantly associated with variation in litter  
183 size among breeds, we retrieved litter size data from 10,810 litters of 224 breeds  
184 registered in the Norwegian Kennel Club (33). For these data, we were able to obtain  
185 average number of pups per litter values for 60 of the 162 dog breeds with overlapping

186 genetic data (Supplementary Table 1). Among these 60 breeds, median litter size is 5.55  
187 pups, with a maximum 8.9 in Rhodesian Ridgebacks and a minimum of 2.4 in Pomeranians  
188 (Supplementary Figure 3B).

189  
190 To identify SNPs, and genes proximal to them, that are significantly associated with the  
191 observed variation in litter size across domestic dog breeds, we conducted an across-breed  
192 GWAS analysis using 115,683 SNPs and litter size data from 60 dog breeds (Figure 1B,  
193 Supplementary Figure 2B). We identified three significant SNPs intersecting three genes,  
194 namely keratin 71 (*KRT71*, uncorrected  $P = 2.2 \times 10^{-8}$ ), RNA Terminal Phosphate  
195 Cyclase-Like 1 (*RCL1*, uncorrected  $P = 2.6 \times 10^{-8}$ ) and microphthalmia-associated  
196 transcription factor (*MITF*, uncorrected  $P = 3.5 \times 10^{-7}$ ). The *KRT71* SNP is the same  
197 variant that associated with variation in cesarean section rate described above. Another  
198 three significant SNPs were found in intergenic regions; two were nearby genes *MSRB3*  
199 (methionine sulfoxide reductase B3, uncorrected  $P = 1.3 \times 10^{-7}$ ) and *MSANTD1*  
200 (Myb/SANT DNA binding domain containing, uncorrected  $P = 1.5 \times 10^{-9}$ ), respectively.  
201 The final variant was near an RNA of unknown function (ENSCAFG00000021196,  
202 uncorrected  $P = 3.8 \times 10^{-10}$ ).

203  
204 The *RCL1* SNP (chromosome 1: 93,189,363) is found in the intron between exons 7 and 8.  
205 *RCL1* functions in the maturation of 18s RNA (51) and is associated with cervical cancer;  
206 one role of the gene in this cancer pathology is thought to involve the regulation of insulin  
207 receptors (51). Additionally, a rare missense variation in *RCL1* was recently associated with  
208 depression (52).

209

210 The *MITF* SNP (chromosome 20: 21,848,176) is found in the intron between exons 4 and 5.  
211 *MITF* plays an integral role in the development of neural crest-derived melanocytes and  
212 optic cup-derived retinal pigment epithelial cells. In human melanocytes, *MITF* is a  
213 regulator of *DIAPH1*, a member of the formin gene family whose members are highly  
214 expressed in reproductive tissues and have been associated with a variety of reproductive  
215 phenotypes (53-57). *DIAPH1* expression is increased in spontaneous term and preterm  
216 labor myometrial tissues (58). In domesticated animals, *MITF* is a well characterized gene  
217 associated with coat color (36,59). In humans, *MITF* is expressed in melanocytes, as well as  
218 reproductive tissues including the endometrium and cervix (Figure 2)(42).

219

220 Another SNP (chromosome 10: 8,114,328) significantly associated with litter size is found  
221 in the intergenic region downstream of *MSRB3*, whose protein product catalyzes the  
222 reduction of methionine-R-sulfoxides to methionine and repairs oxidatively damaged  
223 proteins (60,61). In humans, mutations in *MSRB3* are associated with deafness (62).  
224 Epigenetic changes of *MSRB3* in the fetus during pregnancy may affect length of gestation,  
225 with increased DNA methylation correlated with increased gestational age (63,64).  
226 Furthermore, *MSRB3* shows an increase in mRNA expression in ripe (at term) versus  
227 unripe human uterine cervix, implying that *MSRB3* functions to ripen the cervix before the  
228 onset of labor (65). In previous morphological studies in dogs, *MSRB3* is associated with  
229 ear erectness (36).

230

231 The last SNP (chromosome 6: 61,062,626) that is significantly associated with litter size is  
232 located downstream of *MSANTD1*, which is part of a gene network believed to aid in cell-to-  
233 cell signaling and interaction, hematological system development and function, and  
234 immune cell trafficking (66). *MSANTD1* has been identified in two independent studies as a  
235 candidate gene for the determination of black coat color in goats (67,68).

236

### 237 **Five genetic loci significantly associate with stillbirth rate**

238 To examine whether there are SNPs that are significantly associated with variation in  
239 stillbirth rate among breeds, we retrieved data for stillbirth rates for 57 of the 162 dog  
240 breeds (Supplementary Table 1). The data covers 10,810 litters of 224 breeds registered in  
241 the Norwegian Kennel Club and defines perinatal mortality as the sum of stillborn puppies  
242 and puppies that died during the first week after birth (69). Among these 57 breeds with  
243 overlapping genomic data, the median stillbirth rate is 4.2 pups, with a maximum rate of  
244 12.3% in Saint Bernards and a minimum of 0% in Basenjis and Italian Greyhounds  
245 (Supplementary Figure 3C).

246

247 To test if any SNPs are significantly associated with the observed variation in stillbirth rate  
248 across domestic dog breeds, we conducted an across-breed GWAS analysis using 115,683  
249 SNPs and stillbirth rate data from 56 dog breeds (Figure 1C, Supplementary Figure 2C). We  
250 identified five significant SNPs; four intersecting 4 genes, namely nuclear protein body  
251 SP140 (*SP140*, uncorrected  $P = 2.8 \times 10^{-8}$ ), 5-Hydroxytryptamine receptor 2C (*HTR2C*,  
252 uncorrected  $P = 2.0 \times 10^{-7}$ ), keratin 71 (*KRT71*, uncorrected  $P = 3.2 \times 10^{-9}$ ), and  
253 microphthalmia-associated transcription factor (*MITF*, uncorrected  $P = 1.4 \times 10^{-7}$ ), and

254 one in an intergenic region near a snoRNA (ENSCAFG00000027305, uncorrected  
255  $P = 1.3 \times 10^{-7}$ ) of unknown function. The *KRT71* SNP associated with variation in  
256 stillbirth rate is the same one as that associated with variation in cesarean section rate and  
257 litter size described above. Similarly, the *MITF* SNP associated with variation in stillbirth  
258 rate is the same as that associated with litter size.

259

260 The *SP140* SNP (chromosome 25: 42,482,266) resides in the intro between exons 4 and 5.  
261 *SP140* is the lymphoid-restricted homolog of *SP100* expressed in mature B cells, as well as  
262 some T cells (70). High levels of *SP140* mRNA are detected in human spleen and peripheral  
263 blood leukocytes, but not other human tissues (Bloch et al., 1996). *SP140* expression has  
264 been implicated in innate response to immunodeficiency virus type 1 (71). Finally, *SP140*  
265 was the gene showing the largest difference in expression level between normal and  
266 preeclamptic placentas (72).

267

268 The *HTR2C* SNP (chromosome X: 87,378,551) is located in the intron between exons 3 and  
269 4. *HTR2C* is one of the most important and extensively studied serotonin receptors (73).  
270 *HTR2C* has ten fixed SNP differences between dogs and wolves, and also belongs to the  
271 behavioral fear response (74). Additionally, *HTR2C* is differentially expressed in the brain  
272 between tame and aggressive mice and foxes (75), providing additional evidence for its  
273 involvement in the tame behaviors of domesticated dogs (74).

274

275 **Four genetic loci significantly associate with gestation length**

276 To examine whether there is variation in gestation length among breeds, we identified  
277 individual gestation length averages by breed predominantly in breeder handbooks.  
278 Utilizing breeders' handbooks, we were able to identify gestation length means for a total  
279 of 23 of the 162 dog breeds that we had genotypic data for (Supplementary Table 1).  
280 Among these 23 breeds, the median gestation length is 62.2 days, with a maximum length  
281 of 65.3 in beagles and a minimum of 60.1 in the Alaskan Malamute (Supplementary Figure  
282 3D).

283

284 To identify SNPs, and genes proximal to them, that are significantly associated with the  
285 observed variation in gestation length across domestic dog breeds, we conducted an  
286 across-breed GWAS analysis using 115,683 SNPs and gestation length data from 23 dog  
287 breeds (Figure 1D, Supplementary Figure 2D). Our analysis identified six significantly  
288 associated SNPs that mapped to 4 genes, namely solute carrier family 9 (*SLC9A8*,  
289 uncorrected  $P = 3.7 \times 10^{-11}$ ), calcium channel, voltage-dependent, alpha-2/delta  
290 Subunit 3 (*CACNA2D3*, uncorrected  $P = 3.1 \times 10^{-7}$ ), microtubule associated tumor  
291 suppressor candidate 2 (*MTUS2*, uncorrected  $P = 3.6 \times 10^{-7}$ ), and helicase family  
292 member 1 (*HFM1*, uncorrected  $P = 4.0 \times 10^{-7}$ ), and two lincRNAs  
293 (ENSCAFG00000037743, uncorrected  $P = 4.4 \times 10^{-7}$ , and ENSCAFG00000039067,  
294 uncorrected  $P = 1.6 \times 10^{-7}$ ) whose function is unknown.

295

296 The first significantly associated SNP (chromosome 24: 36,399,705) resides in intron 78 of  
297 *SLC9A8*, an integral transmembrane protein that exchanges extracellular Na<sup>+</sup> for  
298 intracellular H<sup>+</sup>. *SLC9A8* serves multiple functions, including intracellular pH homeostasis,

299 cell volume regulation, and electroneutral NaCl absorption in epithelia (76). Knockout male  
300 mice have impaired luteinizing hormone-stimulated cAMP production and are infertile,  
301 despite normal morphology of their reproductive system and normal behavior (77).  
302 *SLC9A8* is expressed ubiquitously (Figure 2) (42), an expression pattern suggestive of  
303 involvement in housekeeping functions.

304

305 The second SNP (chromosome 20: 35,206,774) is found in the intron between exons 26 and  
306 27 of *CACNA2D3*. This gene is one of four members of the alpha-2/delta subunit three  
307 family of the voltage-dependent calcium (Ca<sup>2+</sup>) channel complex, regulating the influx of  
308 Ca<sup>2+</sup> ions entering the cell upon membrane polarization (78). The regulation of calcium is a  
309 fundamental process relevant to life at fertilization, and subsequent control of  
310 development and differentiation of cells (79). In previous studies in humans, *CACNA2D3* is  
311 differentially methylated in the amnion between normal and preeclamptic pregnancies  
312 (80) and in blood between extreme preterm and term infants at birth (55,81). Additionally,  
313 *CACNA2D3* is one of four genes recently described as influencing cranial morphology in  
314 human populations (82). In other domesticated animals, *CACNA2D3* is downregulated by  
315 Colony Stimulating Factor 2 (*CSF2*) in the trophoectoderm of pregnant cattle, which  
316 increases the ability of the preimplantation embryo to advance to the blastocyst stage (83).  
317 In the closely related wolf, *CACNA2D3* is under diversifying selection associated with  
318 environmental adaptations to altitude (84-86).

319

320 The third significantly associated SNP (chromosome 25: 10,481,606) falls in a large  
321 intronic region of the *MTUS2* gene. The protein product of *MTUS2* is cardiac zipper protein

322 (CAZIP), a member of a class of proteins that interact with angiotensin II receptor  
323 interacting proteins (ATIP) (87). *MTUS2* plays a role in the development and function of the  
324 heart and nervous system in vertebrates (88).

325

326 The fourth SNP (chromosome 6: 57,457,184) is located in the 3' intron of *HFM1*, a DNA  
327 helicase that confers genome integrity in germline tissues (89). *HMF1* plays a role in  
328 meiotic recombination implying a major evolutionary role through the creation of diverse  
329 offspring. In mice, deletion of *HFM1* eliminates a major fraction of cross over events (90),  
330 whereas in cattle *HMF1* is associated with both fertility and milk production in Holstein  
331 cattle (91), as well as with alteration of global recombination rates in Holstein, Holstein-  
332 Friesian, Jersey, and crossbred individuals (92).

333

## 334 **Discussion**

335 Mammals exhibit a great deal of variation in their reproductive traits, yet remarkably little  
336 is known about the genetic basis of these traits. To begin to address this, we used GWAS  
337 analyses to examine the genetic basis of four reproductive traits (cesarean section rate,  
338 stillbirth rate, litter size, and gestation length) across up to 97 domestic dog breeds. We  
339 identified several significant genetic associations for each trait (Figure 1).

340

341 Six of the 14 genetic variants that we found to be associated with reproductive trait  
342 variation have been previously identified to be involved in diverse traits associated with  
343 dog domestication (Table 2), such as brachycephaly and coat curl and color, suggesting that  
344 selection for signature traits of dog breeds may have also directly or indirectly influenced



345 variation in reproductive traits. For example, one of the variants that we found to be  
346 associated with cesarean section rate is in an intron of *SMOC2*, a gene previously associated  
347 with brachycephaly in dogs (40,41). Brachycephaly, the shortening and widening of the  
348 muzzle and skull, is present in several “fighting” breeds such as Boxer, Boston Terrier, and  
349 Bulldog, and is thought to have been originally artificially selected on the basis that a  
350 shorter and wider cranial shape would enhance the dog’s biting power (93). Interestingly,  
351 one of the traits that associated with brachycephaly is cephalopelvic disproportion (94), a  
352 significant medical condition that can result in the death of both the litter and the bitch due  
353 to the inability of the pups to pass through the pelvic canal. The negative effects of  
354 cephalopelvic disproportion are alleviated by cesarean section, which not only allows these  
355 breeds to reproduce but also enables the continued application of artificial selection for the  
356 most extreme cranial morphology (40). Whether the *SMOC2* variant identified directly  
357 influences parturition and birth timing in dogs (in humans, *SMOC2* is highly expressed in  
358 several reproductive tissues; see Figure 2 and Ref. (42) or indirectly leads to adverse  
359 pregnancy outcomes (e.g. brachycephalic cranial morphology leading to cesarean section)  
360 remains unknown. It is highly likely, however, that the association between *SMOC2* and  
361 brachycephaly came first, paving the way for the subsequent association of both with  
362 cesarean section rate.

363  
364 Several of the significantly associated genes that we identified in dogs appear to also be  
365 associated with reproductive phenotypes in humans. This suggests the possibility that the  
366 artificial selection that gave rise to dog breeds may have also contributed to the observed  
367 variation in their reproductive traits. For example, a member of the gene family for a

368 subunit of the voltage-dependent calcium channel complex, *CACNA2D3*, which is associated  
369 with gestation length in our study, has been shown to be both differentially methylated in  
370 amnion between normal and preeclamptic human pregnancies (80), and in blood between  
371 extreme preterm and term infants at birth (55,81). Furthermore, expression of *MSRB3*,  
372 which is associated with litter size in our study, is elevated in ripe (at term) versus unripe  
373 human uterine cervix and may be involved in the onset of labor (65). Finally, a few of the  
374 other genes significantly associated with reproductive traits (*SMOC2* and *MITF*) are also  
375 known to be expressed in human reproductive tissues (42)(Figure 2).

376

## 377 **Methods**

378 **Genotypic and Phenotypic Data.** To identify SNPs that are significantly associated with  
379 reproductive traits, we used a previously published data set containing 160,727 SNPs from  
380 4,342 individual dogs across 162 breeds genotyped using the Illumina 173k CanineHD  
381 array that were downloaded from  
382 <http://datadryad.org/resource/doi:10.5061/dryad.266k4> (35,95). Following the original  
383 authors, SNPs with a genotyping rate (i.e., the proportion of genotypes per marker with  
384 non-missing data) below 95% and heterozygosity ratios (i.e., the ratio of the number of  
385 heterozygous SNPs divided by the number of non-reference homozygous SNPs) below 0.25  
386 or above 1.0 were removed.

387

388 Phenotypic reproductive trait data for litter size (number of pups), cesarean rate, stillbirth  
389 rate, and gestation length across 128 breeds were collected from a variety of breeder's  
390 handbook and primary journal articles (33,69,94,96-104) (see also Supplementary

391 Material 1). We also included body mass as a control trait. Each breed was assigned the  
392 average breed value for each phenotype; the full list of the values for all four reproductive  
393 traits and body mass across the 128 breeds is provided in supplementary table 1. For the  
394 body mass control, our collected trait values overlapped with the genotypic data (35,95)  
395 for 101 breeds corresponding to 3,384 individuals (Table 1). For the reproductive traits,  
396 our collected cesarean section rate trait values overlapped with the genotypic data for 95  
397 breeds (3,194 individuals), our litter size trait values for 60 breeds (2,617 individuals), our  
398 stillbirth rate values for 56 breeds (2,590 individuals), and our gestation length values for  
399 23 breeds (1,908 individuals) (Table 1).

400

401 **Genome Wide Association (GWAS) Analyses.** To test SNPs for associations with the four  
402 reproductive traits of interest, we conducted a GWAS analysis for each individual trait  
403 using body mass as a covariate, and accounting for kinship, as well as for body mass as a  
404 proof of concept. All GWAS analyses were run using a linear-mixed model as implemented  
405 in the program GEMMA, version 0.94 (105). Numerous studies have shown that the vast  
406 majority of morphological, ecological and physiological traits vary as a function of an  
407 organism's body mass (106-108) as well as a function of kinship (35,36). Most notably for  
408 the purpose of this study, body mass has been previously shown to be strongly correlated  
409 with litter weight (109-111), neonate weight (109-112), and gestation length  
410 (6,109,110,113,114).

411

412 To ensure our analysis reflected the reproductive trait of interest and not SNPs associated  
413 with body mass, we used body mass as a covariate for all reproductive trait analyses. To be

414 able to do so, we pruned our genotypic data so that they included only dog breeds (and  
415 individuals) for which we had both body mass and reproductive trait of interest values (see  
416 Supplementary Table 1).

417

418 To account for population stratification, we calculated a kinship matrix of the included  
419 breeds using GEMMA and included it as a random effect in each association analysis. Each  
420 value of a kinship matrix describes the probability that a particular allele from two  
421 randomly chosen individuals at a given locus is identical by descent (115). Finally, to  
422 control for inflated  $P$  value significance from the testing of multiple hypotheses, we used a  
423 significance threshold of  $P = 4.3 \times 10^{-7}$  (Bonferroni cutoff of  $\alpha = 0.05$ ,  $N = 115,574$ ) for  
424 all analyses. All reported  $P$  values are uncorrected Wald's  $P$  values as calculated in PLINK.

425

426 Finally, to reduce potential error stemming from SNP misidentification in our analyses, we  
427 included only SNPs with a minor allele frequency (MAF)  $> 0.05$ , since SNPs with very low  
428 minor allele frequencies are more prone to error due to the small number of samples that  
429 have the called nucleotide. Furthermore, we pruned SNPs not in complete or near-complete  
430 linkage disequilibrium using a variance inflation factor of 10, using the PLINK command --  
431 indep 100 10 10 (116).

432

433 To gain insight into the genetic elements putatively involved with the traits of interest, we  
434 mapped all SNPs found to be significantly associated with each trait of interest using  
435 custom perl and R scripts to the CanFam3.1.87 dog genome assembly (117,118). Transcript  
436 IDs were mapped to gene names using bioconductor biomaRt interface to the ENSEMBL

437 biomaRt (119). If the significant SNP was outside gene boundaries, we reported the nearest  
438 upstream or downstream gene. Manhattan plots and quantile-quantile plots were  
439 generated using R 3.1.2 (120) with the qqman package (121). Calculation of the  $\lambda$  inflation  
440 parameter, a metric of any existing systematic bias in the data set, was calculated using the  
441 GenABEL R package (122) and was used to interpret Type I error rate in the multiple  
442 testing of GWAS analyses (123).

443

#### 444 **Acknowledgements**

445 We thank members of the Rokas lab and members of the March of Dimes Prematurity  
446 Research Center Ohio Collaborative for useful discussions of this work.

447

#### 448 **Funding Statement**

449 This research was supported by the March of Dimes through the March of Dimes  
450 Prematurity Research Center Ohio Collaborative (to A.R., P.A. and J.A.C) and by the  
451 Burroughs Wellcome Fund (to J.A.C and A.R.). MLJ received partial support through a  
452 Littlejohn Summer Research Scholarship at Vanderbilt University. The funders had no role  
453 in study design, data collection and analysis, decision to publish, or preparation of the  
454 manuscript.

455 **TABLES**

456 **Table 1.** Numbers of breeds and individuals with overlapping phenotypes and genotypes

457 included in our analysis.

| Trait                 | Number of Overlapping Breeds | Number of Overlapping Individuals |
|-----------------------|------------------------------|-----------------------------------|
| Body Mass             | 101                          | 3,384                             |
| Cesarean Section Rate | 97                           | 3,194                             |
| Litter Size           | 60                           | 2,617                             |
| Stillbirth Rate       | 57                           | 2,590                             |
| Gestation Length      | 23                           | 1,908                             |

458

459

460 **Table 2.** Summary of genes that contain or are adjacent to the SNPs that are significantly  
 461 associated with variation in reproductive traits across dog breeds.

| Gene ID      | Gene Name                               | Chr. | rs Number  | Variant                         | Reproductive Trait(s)                               | Domestication-related Trait(s) |
|--------------|---|------|------------|---------------------------------|---|--------------------------------|
| <i>SMOC2</i> | SPARC related modular calcium binding 2 | 1    | rs21966904 | Non-coding (Intron 13)          | Cesarean section rate                               | Brachycephaly (Dogs)           |
| <i>PALM3</i> | paralemmin 3                            | 20   | rs22853767 | Non-coding (3' UTR)             | Cesarean section rate                               | -                              |
| <i>KRT71</i> | keratin                                 | 27   | rs23373415 | Coding (exon 2)                 | Cesarean section rate, litter size, stillbirth rate | Coat phenotypes (Dogs)         |
| <i>CD36</i>  | CD36 glycoprotein                       | 18   | rs22664051 | Intergenic variant (downstream) | Cesarean section rate                               | -                              |
| <i>RCL1</i>  | RNA terminal phosphate cyclase          | 1    | rs21894066 | Non-coding (Intron 7)           | Litter size   | -                              |

|                 |   |    |            |                                 |                              |                              |
|-----------------|---|----|------------|---------------------------------|------------------------------|------------------------------|
|                 | like 1  |    |            |                                 |                              |                              |
| <i>MITF</i>     | melanogenesis associated transcription factor | 20 | rs20848176 | Coding (exon 5)                 | Litter size, stillbirth rate | Coat color (Dogs)            |
| <i>MSRB3</i>    | methionine sulfoxide reductase B3             | 10 | rs22060533 | Intergenic variant (downstream) | Litter size                  | Ear erectness (Dogs)         |
| <i>MSANTD1</i>  | Myb/SANT DNA binding domain containing        | 6  | rs9084938  | Intergenic variant (downstream) | Litter size                  | Black coat color (Goats)     |
| <i>SP140</i>    | nuclear protein body SP140                    | 25 | rs8856304  | Non-coding (intron 4)           | Stillbirth rate              | -                            |
| <i>HTR2C</i>    | 5-hydroxytryptamine receptor 2                | X  | rs24622199 | Non-Coding (intron 2)           | Stillbirth rate              | Tameness (Dogs, Foxes, Mice) |
| <i>SLC9A8</i>   | solute carrier family 9 member A8             | 24 | rs23219089 | Non-coding (intron 7)           | Gestation length             | -                            |
| <i>CACNA2D3</i> | calcium voltage-                              | 20 | rs22853845 | Non-coding                      | Gestation                    | Blastocyst                   |



|              |  |    |            |                          |                     |  |
|--------------|--|----|------------|--------------------------|---------------------|--|
|              | gated channel<br>auxiliary subunit<br>alpha2delta 3          |    |            | (intron 9)               | length              | development<br>(Cattle)                      |
| <i>MTUS2</i> | microtubule<br>associated tumor<br>suppressor<br>candidate 2 | 25 | 10481606   | Non-coding<br>(intron 6) | Gestation<br>length | -  |
| <i>HFM1</i>  | ATP dependent<br>DNA helicase<br>homolog                     | 6  | rs24306896 | Non-coding<br>(intron 4) | Gestation<br>length | Fertility and<br>milk production<br>(Cattle) |

462

463

464

465

466

467 **Figure Legends**

468

469 **Figure 1. Significant GWAS results for reproductive traits in domestic dogs.**

470 Manhattan plots showing the statistical significance of each SNP as a function of genomic  
471 position for (A) cesarean section rate (n = 3,194 individuals, n = 97 breeds), (B) litter size  
472 (n = 2,617 individuals, n = 60 breeds), (C) stillbirth (n = 2,590 individuals, n = 57 breeds),  
473 and (D) gestation length (n = 1,908 individuals, n = 23 breeds). Horizontal line indicates the  
474 significance threshold at  $P = 4.3 \times 10^{-7}$ . Significant SNPs are labels with the intersecting  
475 or nearest gene. Plots were generated in R using the qqman package.

476

477 **Figure 2. Gene expression in human female reproductive tissues of genes that**  
478 **contain or are adjacent to SNPs significantly associated with reproductive traits in**  
479 **domestic dogs.** Raw data were obtained from the Human Protein Atlas database (42).

480

481 **Supplementary Material**

482

483 **Supplementary Figure 1. Recapitulation of SNPs associated with body mass in 101**

484 **domesticated dog breeds.** (A) Body mass distribution for 101 breeds. (B) Manhattan plots

485 showing the statistical significance of each SNP as a function of genomic position for body

486 mass. Plot generated in R using the qqman package. (C) Quantile-quantile plot showing the

487 effectiveness of the stratification correction ( $\lambda = 1.17$ ). Plot generated in R; inflation factor

488 was calculated using the GenABEL package implemented in R.

489

490 **Supplementary Figure 2. Distribution of phenotypic values of the four reproductive**

491 **traits examined in this study across dog breeds.** (A) cesarean section rate ( $n = 97$

492 breeds), (B) litter size ( $n = 60$  breeds), (C) stillbirth rate ( $n = 57$  breeds), and (D) gestation

493 length ( $n = 23$  breeds). Plots were generated in R using the ggplot2 package.

494

495 **Supplementary Figure 3. Quantile-quantile plots for the GWAS analyses of the four**

496 **reproductive traits.** The range for the inflation factor ( $\lambda$ ) for all GWAS analyses is between

497 1.05 – 1.09, indicating the effectiveness of the stratification correction. (A) cesarean section

498 rate ( $\lambda = 1.05$ ), (B) litter size ( $\lambda = 1.05$ ), (C) stillbirth rate ( $\lambda = 1.05$ ), and (D) gestation

499 length ( $\lambda = 1.09$ ). Plots generated in R, and inflation factors were calculated using the

500 GenABEL package implanted in R.

501

502 **Supplementary Table 1. Summary of raw phenotypes for breeds included in analysis.**

503

504 **Supplementary Table 2. Summary of top 50 SNPs associated with body mass.**

505

506 **Supplementary Table 3. Summary of top 50 SNPs associated with cesarean section**  
507 **rate.**

508

509 **Supplementary Table 4. Summary of top 50 SNPs associated with litter size.**

510

511 **Supplementary Table 5. Summary of top 50 SNPs associated with stillbirth rate.**

512

513 **Supplementary Table 6. Summary of top 50 SNPs associated with gestation length.**

514

## 515 References

- 516 1. Derrickson EM. Comparative reproductive strategies of altricial and precocial  
517 eutherian mammals. *Funct Ecol.* 1992;6(1):57–65.
- 518 2. Harrison RM. *Reproduction in Mammals: General Overview*. Chichester, UK: John  
519 Wiley & Sons, Ltd; 2001.
- 520 3. Behringer RR, Eakin GS, Renfree MB. Mammalian diversity: gametes, embryos and  
521 reproduction. *Reprod Fertil Dev.* 2006;18(2):99–9.
- 522 4. Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, et al. PanTHERIA: a  
523 species-level database of life history, ecology, and geography of extant and recently  
524 extinct mammals. *Ecology*. Ecological Society of America; 2009;90(9):2648.
- 525 5. Kiltie RA. Intraspecific Variation in the Mammalian Gestation Period. *Journal of*  
526 *Mammalogy.* 1982;63(4):646–52.
- 527 6. Martin RD, MacLarnon AM. Gestation Period, Neonatal Size and Maternal  
528 Investment in Placental Mammals. *Nature.* 1985;313(5999):220–3.
- 529 7. Cross JC, Werb Z, Fisher SJ. Implantation and the placenta: key pieces of the  
530 development puzzle. *Science.* 1994 Dec 2;266(5190):1508–18.
- 531 8. Enders AC, Carter AM. What Can Comparative Studies of Placental Structure Tell  
532 Us?—A Review. *Placenta.* 2004 Apr;25:S3–S9.
- 533 9. Elliot MG, Crespi BJ. Phylogenetic Evidence for Early Hemochorial Placentation in  
534 Eutheria. *Placenta.* Elsevier Ltd; 2009 Nov 1;30(11):949–67.
- 535 10. Pond CM. The Significance of Lactation in the Evolution of Mammals. *Evolution.*  
536 1977 Mar;31(1):177–99.
- 537 11. Lefèvre CM, Sharp JA, Nicholas KR. Evolution of Lactation: Ancient Origin and  
538 Extreme Adaptations of the Lactation System. [http://dxdoi.org/10.1146/annurev-](http://dxdoi.org/10.1146/annurev-genom-082509-141806)  
539 [genom-082509-141806](http://dxdoi.org/10.1146/annurev-genom-082509-141806). *Annual Reviews*; 2010 Sep 7;11(1):219–38.
- 540 12. Cross JF, Martin RD. Calculation of gestation period and other reproductive  
541 parameters for primates. *Dodo: Journal of the Jersey Wildlife Preservation Trust.*  
542 *Durrell Wildlife Conservation Trust*; 1981;18:30–43.
- 543 13. Hudson R, Cruz Y, Lucio A, Ninomiya J, Martínez-Gómez M. Temporal and  
544 behavioral patterning of parturition in rabbits and rats. *Physiol Behav.* 1999  
545 Jun;66(4):599–604.
- 546 14. Burriss MJ, Blunn CT. Some factors affecting gestation length and birth weight of  
547 beef cattle. *J Anim Sci.* American Society of Animal Science; 1952;11(1):34–41.

- 548 15. Davies Morel MCG, Newcombe JR, Holland SJ. Factors affecting gestation length in  
549 the Thoroughbred mare. *Animal Reproduction Science*. 2002 Dec 16;74(3-4):175-  
550 85.
- 551 16. Dufour DL, Sauter ML. Comparative and evolutionary dimensions of the energetics  
552 of human pregnancy and lactation. *Am J Hum Biol*. Wiley-Blackwell; 2002 Aug  
553 21;14(5):584-602.
- 554 17. Carneiro M, Piorno V, Rubin C-J, Alves JM, Ferrand N, Alves PC, et al. Candidate  
555 genes underlying heritable differences in reproductive seasonality between wild  
556 and domestic rabbits. *Animal Genetics*. 2015 May 22;46(4):418-25.
- 557 18. Sironen A, Uimari P, Nagy S, Paku S, Andersson M, Vilkki J. Knobbed acrosome  
558 defect is associated with a region containing the genes *STK17b* and *HECW2* on  
559 porcine chromosome 15. *BMC Genomics*. BioMed Central; 2010 Dec 9;11(1):699.
- 560 19. Bolormaa S, Pryce JE, Hayes BJ, Goddard ME. Multivariate analysis of a genome-  
561 wide association study in dairy cattle. *Journal of Dairy Science*. Elsevier; 2010 Aug  
562 1;93(8):3818-33.
- 563 20. Maltecca C, Gray KA, Weigel KA, Cassady JP, Ashwell M. A genome-wide association  
564 study of direct gestation length in US Holstein and Italian Brown populations.  
565 *Animal Genetics*. 2011 May 6;42(6):585-91.
- 566 21. Walker R, Gurven M, Hill K, Migliano A, Chagnon N, De Souza R, et al. Growth rates  
567 and life histories in twenty-two small-scale societies. *Am J Hum Biol*.  
568 2006;18(3):295-311.
- 569 22. Horikoshi M, Beaumont RN, Day FR, Warrington NM, Kooijman MN, Fernandez-  
570 Tajes J, et al. Genome-wide associations for birth weight and correlations with adult  
571 disease. *Nature*. 2016 Sep 28;538(7624):248-52.
- 572 23. Plunkett J, Doniger S, Orabona G, Morgan T, Haataja R, Hallman M, et al. An  
573 Evolutionary Genomic Approach to Identify Genes Involved in Human Birth Timing.  
574 Barsh GS, editor. *PLoS Genetics*. 2011 Apr 14;7(4):e1001365.
- 575 24. Zhang H, Baldwin DA, Bukowski RK, Parry S, Xu Y, Song C, et al. A genome-wide  
576 association study of early spontaneous preterm delivery. *Genet Epidemiol*. 2015  
577 Mar;39(3):217-26.
- 578 25. Zhang G, Feenstra B, Bacelis J, Liu X, Muglia LM, Juodakis J, et al. Genetic  
579 Associations with Gestational Duration and Spontaneous Preterm Birth. *N Engl J*  
580 *Med*. 2017 Sep 21;377(12):1156-67.
- 581 26. Cassady JP, Johnson RK, Pomp D, Rohrer GA, Van Vleck LD, Spiegel EK, et al.  
582 Identification of quantitative trait loci affecting reproduction in pigs. *J Anim Sci*.  
583 2001 Mar;79(3):623-33.

- 584 27. Langlois B, Blouin C, Chaffaux S. Analysis of several factors of variation of gestation  
585 loss in breeding mares. *Animal*. 2012 Aug 9;6(12):1925–30.
- 586 28. Johanson JM, Berger PJ, Tsuruta S, Misztal I. A Bayesian threshold-linear model  
587 evaluation of perinatal mortality, dystocia, birth weight, and gestation length in a  
588 Holstein herd. *Journal of Dairy Science*. 2011 Jan;94(1):450–60.
- 589 29. Bagnicka E, Wallin E, Łukaszewicz M, Ådnøy T. Heritability for reproduction traits  
590 in Polish and Norwegian populations of dairy goat. *Small Ruminant Research*. 2007  
591 Apr;68(3):256–62.
- 592 30. Neff MW, Rine J. A Fetching Model Organism. *Cell*. 2006 Jan;124(2):229–31.
- 593 31. Beale H, Ostrander EA. Sizing up dogs. *Current Biology*. Elsevier; 2012 May  
594 8;22(9):R315–6.
- 595 32. Marsden CD, Ortega-Del Vecchyo D, O'Brien DP, Taylor JF, Ramirez O, Vilà C, et al.  
596 Bottlenecks and selective sweeps during domestication have increased deleterious  
597 genetic variation in dogs. *Proc Natl Acad Sci USA*. 2016 Jan 5;113(1):152–7.
- 598 33. Borge KS, Tønnessen R, Nødtvedt A, Indrebø A. Litter size at birth in purebred dogs-  
599 -a retrospective study of 224 breeds. *Theriogenology*. 2011 Mar 15;75(5):911–9.
- 600 34. Evans KM, Adams VJ. Proportion of litters of purebred dogs born by caesarean  
601 section. *J Small Anim Pract*. Blackwell Publishing Ltd; 2010 Feb;51(2):113–8.
- 602 35. Hayward JJ, Castelhana MG, Oliveira KC, Corey E, Balkman C, Baxter TL, et al.  
603 Complex disease and phenotype mapping in the domestic dog. *Nature*  
604 *Communications*. Nature Publishing Group; 1AD;7:1–11.
- 605 36. Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, Lohmueller KE, et al. A  
606 Simple Genetic Architecture Underlies Morphological Variation in Dogs. Hoekstra  
607 HE, editor. *Plos Biol*. Public Library of Science; 2010 Aug 10;8(8):e1000451.
- 608 37. Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Pielberg GR, Sigurdsson S, et al.  
609 Identification of Genomic Regions Associated with Phenotypic Variation between  
610 Dog Breeds using Selection Mapping. Akey JM, editor. *PLoS Genetics*. Public Library  
611 of Science; 2011 Oct 13;7(10):e1002316.
- 612 38. Marchant TW, Johnson EJ, McTeir L, Johnson CI, Gow A, Liuti T, et al. Canine  
613 Brachycephaly Is Associated with a Retrotransposon-Mediated Missplicing of  
614 SMO2. *Curr Biol*. Elsevier; 2017 Jun 5;27(11):1573–6.
- 615 39. Karlsson EK, Lindblad-Toh K. Leader of the pack: gene mapping in dogs and other  
616 model organisms. Nature Publishing Group. Nature Publishing Group; 2008  
617 Sep;9(9):713–25.

- 618 40. Bannasch D, Young A, Myers J, Truvé K, Dickinson P, Gregg J, et al. Localization of  
619 Canine Brachycephaly Using an Across Breed Mapping Approach. Fairhead C,  
620 editor. PLoS ONE. Public Library of Science; 2010 Mar 10;5(3):e9632.
- 621 41. Marchant TW, Johnson EJ, McTeir L, Johnson CI, Gow A, Liuti T, et al. Canine  
622 Brachycephaly Is Associated with a Retrotransposon-Mediated Missplicing of  
623 SMOC2. *Current Biology*. Elsevier Ltd; 2017 May 22;:1–35.
- 624 42. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al.  
625 Tissue-based map of the human proteome. *Science*. American Association for the  
626 Advancement of Science; 2015 Jan 23;347(6220):1260419–9.
- 627 43. Rose AB. Intron-mediated regulation of gene expression. *Curr Top Microbiol*  
628 *Immunol*. 2008;326:277–90.
- 629 44. Basile M, Lin R, Kabbani N, Karpa K, Kilimann M, Simpson I, et al. Paralemmin  
630 interacts with D3 dopamine receptors: implications for membrane localization and  
631 cAMP signaling. *Arch Biochem Biophys*. 2006 Feb 1;446(1):60–8.
- 632 45. Kutzleb C, Sanders G, Yamamoto R, Wang X, Lichte B, Petrasch-Parwez E, et al.  
633 Paralemmin, a prenyl-palmitoyl-anchored phosphoprotein abundant in neurons  
634 and implicated in plasma membrane dynamics and cell process formation. *J Cell*  
635 *Biol*. The Rockefeller University Press; 1998 Nov 2;143(3):795–813.
- 636 46. Chen X, Wu X, Zhao Y, Wang G, Feng J, Li Q, et al. A novel binding protein of single  
637 immunoglobulin IL-1 receptor-related molecule: Paralemmin-3. *Biochem Biophys*  
638 *Res Commun*. 2011 Jan 28;404(4):1029–33.
- 639 47. Langbein L, Rogers MA, Praetzel S, Winter H, Schweizer J. K6irs1, K6irs2, K6irs3,  
640 and K6irs4 Represent the Inner-Root-Sheath-Specific Type II Epithelial Keratins of  
641 the Human Hair Follicle1. *J Invest Dermatol*. Elsevier Masson SAS; 2003 Jan  
642 1;120(4):512–22.
- 643 48. Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, et al. Coat variation in  
644 the domestic dog is governed by variants in three genes. *Science*. American  
645 Association for the Advancement of Science; 2009 Oct 2;326(5949):150–3.
- 646 49. Tandon NN, Kralisz U, Jamieson GA. Identification of glycoprotein IV (CD36) as a  
647 primary receptor for platelet-collagen adhesion. *J Biol Chem*. 1989 May  
648 5;264(13):7576–83.
- 649 50. Febbraio M, Hajjar DP, Silverstein RL. CD36: a class B scavenger receptor involved  
650 in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *J Clin Invest*.  
651 American Society for Clinical Investigation; 2001 Sep;108(6):785–91.
- 652 51. Lyng H, Brøvig RS, Svendsrud DH, Holm R, Kaalhus O, Knutstad K, et al. Gene  
653 expressions and copy numbers associated with metastatic phenotypes of uterine



- 654 cervical cancer. *BMC Genomics*. BioMed Central; 2006 Oct 20;7(1):268.
- 655 52. Amin N, de Vrij FMS, Baghdadi M, Brouwer RWW, van Rooij JGJ, Jovanova O, et al. A  
656 rare missense variant in RCL1 segregates with depression in extended families.  
657 *Molecular Psychiatry*. Nature Publishing Group; 2017 Mar 21;58:1323.
- 658 53. Carreira S, Goodall J, Denat L, Rodriguez M, Nuciforo P, Hoek KS, et al. Mitf  
659 regulation of Dia1 controls melanoma proliferation and invasiveness. *Genes Dev*.  
660 2006 Dec 15;20(24):3426–39.
- 661 54. Lamm KYB, Johnson ML, Baker Phillips J, Muntifering MB, James JM, Jones HN, et al.  
662 Inverted formin 2 regulates intracellular trafficking, placentation, and pregnancy  
663 outcome. *Elife*. eLife Sciences Publications Limited; 2018 Jan 8;7:742.
- 664 55. Cruickshank MN, Oshlack A, Theda C, Davis PG, Martino D, Sheehan P, et al. Analysis  
665 of epigenetic changes in survivors of preterm birth reveals the effect of gestational  
666 age and evidence for a long term legacy. *Genome Med*. BioMed Central;  
667 2013;5(10):96.
- 668 56. Elovitz MA, Brown AG, Anton L, Gilstrap M, Heiser L, Bastek J. Distinct cervical  
669 microRNA profiles are present in women destined to have a preterm birth. *Am J*  
670 *Obstet Gynecol*. Elsevier; 2014 Mar 1;210(3):221.e1–221.e11.
- 671 57. Montenegro D, Romero R, Kim SS, Tarca AL, Draghici S, Kusanovic JP, et al.  
672 Expression patterns of microRNAs in the chorioamniotic membranes: a role for  
673 microRNAs in human pregnancy and parturition. *J Pathol*. John Wiley & Sons, Ltd;  
674 2009 Jan;217(1):113–21.
- 675 58. Lartey J, Smith M, Pawade J, Strachan B, Mellor H, Bernal AL. Up-Regulation of  
676 Myometrial RHO Effector Proteins (PKN1 and DIAPH1) and CPI-17 (PPP1R14A)  
677 Phosphorylation in Human Pregnancy Is Associated with Increased GTP-RHOA in  
678 Spontaneous Preterm Labor. *Biol Reprod*. Oxford University Press; 2007 Jun  
679 1;76(6):971–82.
- 680 59. Wang X, Liu J, Zhou G, Guo J, Yan H, Niu Y, et al. Whole-genome sequencing of eight  
681 goat populations for the detection of selection signatures underlying production  
682 and adaptive traits. *Sci Rep*. Nature Publishing Group; 2016 Dec 9;:1–10.
- 683 60. Brot N, Weissbach L, Werth J, Weissbach H. Enzymatic reduction of protein-bound  
684 methionine sulfoxide. *Proc Natl Acad Sci USA*. National Academy of Sciences; 1981  
685 Apr;78(4):2155–8.
- 686 61. Kim H-Y, Gladyshev VN. Characterization of mouse endoplasmic reticulum  
687 methionine-R-sulfoxide reductase. *Biochem Biophys Res Commun*. 2004 Aug  
688 6;320(4):1277–83.
- 689 62. Ahmed ZM, Yousaf R, Lee BC, Khan SN, Lee S, Lee K, et al. Functional null mutations

- 690 of MSRB3 encoding methionine sulfoxide reductase are associated with human  
691 deafness DFNB74. *Am J Hum Genet.* 2011 Jan 7;88(1):19–29.
- 692 63. Lee H, Jaffe AE, Feinberg JL, Tryggvadottir R, Brown S, Montano C, et al. DNA  
693 methylation shows genome-wide association of NFIX, RAPGEF2 and MSRB3 with  
694 gestational age at birth. *International Journal of Epidemiology.* 2012  
695 Feb;41(1):188–99.
- 696 64. Schierding W, O'Sullivan JM, Derraik JGB, Cutfield WS. Genes and post-term birth:  
697 late for delivery. *BMC Res Notes. BioMed Central;* 2014 Oct 14;7(1):720.
- 698 65. Hassan SS, Romero R, Tarca AL, Nhan-Chang C-L, Vaisbuch E, Erez O, et al. The  
699 transcriptome of cervical ripening in human pregnancy before the onset of labor at  
700 term: identification of novel molecular functions involved in this process. *J Matern  
701 Fetal Neonatal Med.* 2009 Dec;22(12):1183–93.
- 702 66. Yu F, Shen X-Y, Fan L, Yu Z-C. Genome-wide analysis of genetic variations assisted  
703 by Ingenuity Pathway Analysis to comprehensively investigate potential genetic  
704 targets associated with the progression of hepatocellular carcinoma. *Eur Rev Med  
705 Pharmacol Sci.* 2014;18(15):2102–8.
- 706 67. Benjelloun B, Alberto FJ, Streeter I, Boyer F, Coissac E, Stucki S, et al. Characterizing  
707 neutral genomic diversity and selection signatures in indigenous populations of  
708 Moroccan goats (*Capra hircus*) using WGS data. *Front Genet. Frontiers Media SA;*  
709 2015;6:4447.
- 710 68. Wang X, Liu J, Zhou G, Guo J, Yan H, Niu Y, et al. Whole-genome sequencing of eight  
711 goat populations for the detection of selection signatures underlying production  
712 and adaptive traits. *Sci Rep. Nature Publishing Group;* 2016 Dec 12;6(1):38932.
- 713 69. Tønnessen R, Borge KS, Nødtvedt A, Indrebø A. Canine perinatal mortality: A cohort  
714 study of 224 breeds. *THE. Elsevier;* 2012 Jun 1;77(9):1788–801.
- 715 70. Bloch DB, la Monte de SM, Guigaouri P, Filippov A, Bloch KD. Identification and  
716 characterization of a leukocyte-specific component of the nuclear body. *J Biol Chem.*  
717 1996 Nov 15;271(46):29198–204.
- 718 71. Madani N, Millette R, Platt EJ, Marin M, Kozak SL, Bloch DB, et al. Implication of the  
719 lymphocyte-specific nuclear body protein Sp140 in an innate response to human  
720 immunodeficiency virus type 1. *J Virol. American Society for Microbiology (ASM);*  
721 2002 Nov;76(21):11133–8.
- 722 72. Heikkilä A, Tuomisto T, Häkkinen S-K, Keski-Nisula L, Heinonen S, Ylä-Herttuala S.  
723 Tumor suppressor and growth regulatory genes are overexpressed in severe early-  
724 onset preeclampsia--an array study on case-specific human preeclamptic placental  
725 tissue. *Acta Obstet Gynecol Scand. Munksgaard International Publishers;* 2005  
726 Jul;84(7):679–89.

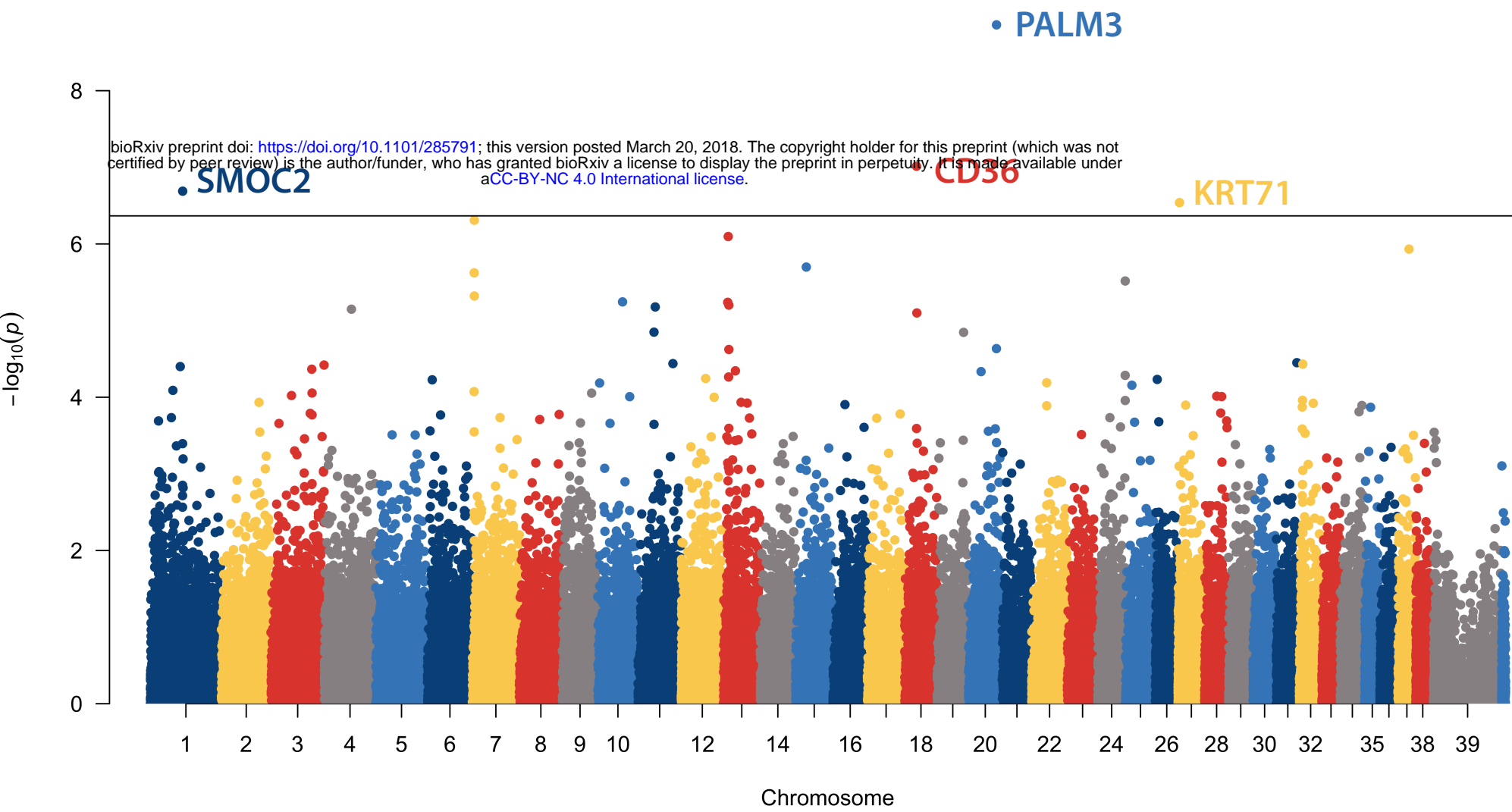
- 727 73. Drago A, Serretti A. Focus on HTR2C: A possible suggestion for genetic studies of  
728 complex disorders. *American Journal of Medical Genetics Part B: Neuropsychiatric*  
729 *Genetics*. Wiley Subscription Services, Inc., A Wiley Company; 2009 Jul  
730 5;150B(5):601–37.
- 731 74. Li Y, vonHoldt BM, Reynolds A, Boyko AR, Wayne RK, Wu D-D, et al. Artificial  
732 selection on brain-expressed genes during the domestication of dog. *Molecular*  
733 *Biology and Evolution*. 2013 Aug;30(8):1867–76.
- 734 75. Kukekova AV, Johnson JL, Teiling C, Li L, Oskina IN, Kharlamova AV, et al. Sequence  
735 comparison of prefrontal cortical brain transcriptome from a tame and an  
736 aggressive silver fox (*Vulpes vulpes*). *BMC Genomics*. Third. BioMed Central; 2011  
737 Oct 3;12(1):482.
- 738 76. Xu H, Chen H, Dong J, Lynch R, Ghishan FK. Gastrointestinal distribution and kinetic  
739 characterization of the sodium-hydrogen exchanger isoform 8 (NHE8). *Cell Physiol*  
740 *Biochem*. Karger Publishers; 2008;21(1-3):109–16.
- 741 77. Xu H, Chen H, Li J, Zhao Y, Ghishan FK. Disruption of NHE8 expression impairs  
742 Leydig cell function in the testes. *American Journal of Physiology- Cell Physiology*.  
743 *American Physiological Society*; 2015 Feb 15;308(4):C330–8.
- 744 78. Jin Y, Cui D, Ren J, Wang K, Zeng T, Gao L. CACNA2D3 is downregulated in gliomas  
745 and functions as a tumor suppressor. *Mol Carcinog*. 2017 Mar;56(3):945–59.
- 746 79. Berridge MJ, Bootman MD, Lipp P. Calcium--a life and death signal. *Nature*. 1998  
747 Oct 15;:645–8.
- 748 80. Suzuki M, Maekawa R, Patterson NE, Reynolds DM, Calder BR, Reznik SE, et al.  
749 Amnion as a surrogate tissue reporter of the effects of maternal preeclampsia on  
750 the fetus. *Clin Epigenetics*. BioMed Central; 2016;8(1):67.
- 751 81. Eidem HR, Ackerman WE, McGary KL, Abbot P, Rokas A. Gestational tissue  
752 transcriptomics in term and preterm human pregnancies: a systematic review and  
753 meta-analysis. *BMC Med Genomics*. BioMed Central; 2015 Jun 5;8(1):27.
- 754 82. Paternoster L, Zhurov AI, Toma AM, Kemp JP, St Pourcain B, Timpson NJ, et al.  
755 Genome-wide association study of three-dimensional facial morphology identifies a  
756 variant in PAX3 associated with nasion position. *Am J Hum Genet*. 2012 Mar  
757 9;90(3):478–85.
- 758 83. Ozawa M, Sakatani M, Dobbs KB, Kannampuzha-Francis J, Hansen PJ. Regulation of  
759 gene expression in the bovine blastocyst by colony stimulating factor 2. *BMC Res*  
760 *Notes*. BioMed Central; 2016 Apr 29;9(1):250.
- 761 84. Pilot M, Greco C, vonHoldt BM, Jędrzejewska B, Randi E, Jędrzejewski W, et al.  
762 Genome-wide signatures of population bottlenecks and diversifying selection in

- 763 European wolves. *Heredity*. Nature Publishing Group; 2014 Apr;112(4):428–42.
- 764 85. Schweizer RM, vonHoldt BM, Harrigan R, Knowles JC, Musiani M, Coltman D, et al.  
765 Genetic subdivision and candidate genes under selection in North American grey  
766 wolves. *Mol Ecol*. 2016 Jan;25(1):380–402.
- 767 86. Zhang W, Fan Z, Han E, Hou R, Zhang L, Galaverni M, et al. Hypoxia adaptations in  
768 the grey wolf (*Canis lupus chanco*) from Qinghai-Tibet Plateau. Akey JM, editor.  
769 *PLoS Genetics*. 2014 Jul;10(7):e1004466.
- 770 87. Rodrigues-Ferreira S, Nahmias C. An ATIPical family of angiotensin II AT2 receptor-  
771 interacting proteins. *Trends Endocrinol Metab*. 2010 Nov;21(11):684–90.
- 772 88. Puy Du L, Beqqali A, Monshouwer-Kloots J, Haagsman HP, Roelen BAJ, Passier R.  
773 CAZIP, a novel protein expressed in the developing heart and nervous system. *Dev*  
774 *Dyn*. Wiley-Liss, Inc; 2009 Nov;238(11):2903–11.
- 775 89. Tanaka K, Miyamoto N, Shouguchi-Miyata J, Ikeda J-E. HFM1, the human homologue  
776 of yeast Mer3, encodes a putative DNA helicase expressed specifically in germ-line  
777 cells. *DNA Sequence*. 2009 Jul 11;17(3):242–6.
- 778 90. Guiraldelli MF, Eyster C, Wilkerson JL, Dresser ME, Pezza RJ. Mouse HFM1/Mer3 is  
779 required for crossover formation and complete synapsis of homologous  
780 chromosomes during meiosis. Hawley RS, editor. *PLoS Genetics*. Public Library of  
781 Science; 2013 Mar;9(3):e1003383.
- 782 91. Pimentel ECG, Bauersachs S, Tietze M, Simianer H, Tetens J, Thaller G, et al.  
783 Exploration of relationships between production and fertility traits in dairy cattle  
784 via association studies of SNPs within candidate genes derived by expression  
785 profiling. *Animal Genetics*. Blackwell Publishing Ltd; 2011 Jun;42(3):251–62.
- 786 92. Kadri NK, Harland C, Faux P, Cambisano N, Karim L, Coppieters W, et al. Coding and  
787 noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect  
788 recombination rate in cattle. *Genome Research*. Cold Spring Harbor Lab; 2016  
789 Oct;26(10):1323–32.
- 790 93. Ellis JL, Thomason J, Kebreab E, Zubair K, France J. Cranial dimensions and forces of  
791 biting in the domestic dog. *J Anat*. Blackwell Publishing Ltd; 2009 Mar  
792 1;214(3):362–73.
- 793 94. Evans KM, Adams VJ. Proportion of litters of purebred dogs born by caesarean  
794 section. *Journal of Small Animal Practice*. 2010 Feb;51(2):113–8.
- 795 95. Hayward JJ, Castelhana MG, Oliveira KC, Corey E, Balkman C, Baxter TL, et al. Data  
796 from: Complex disease and phenotype mapping in the domestic dog. *Dryad Data*  
797 *Repository* [Internet]. Available from: <http://dx.doi.org/10.5061/dryad.266k4>

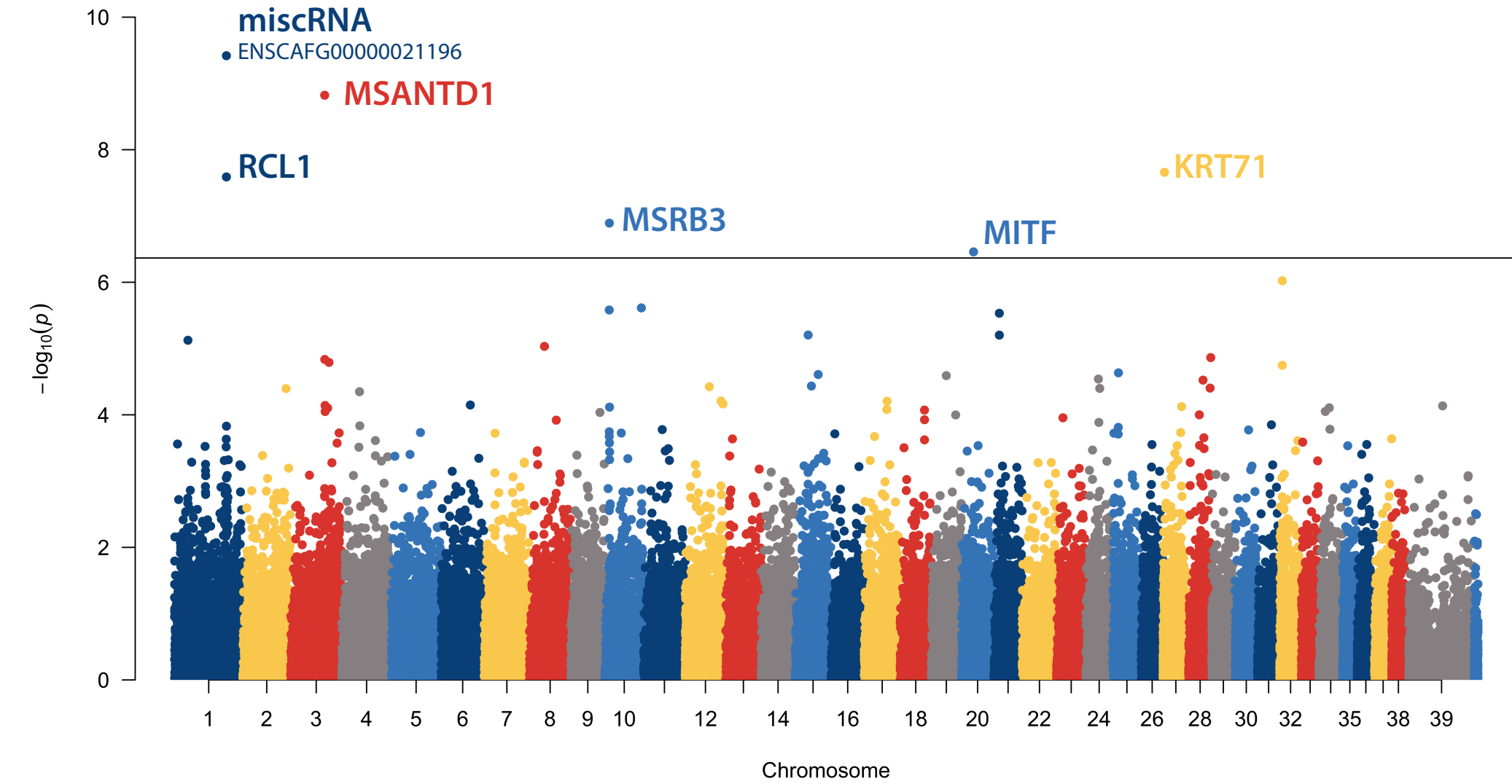
- 798 96. Chatdarong K, Tummaruk P, Sirivaidyapong S, Raksil S. Seasonal and breed effects  
799 on reproductive parameters in bitches in the tropics: a retrospective study. *J Small*  
800 *Anim Pract.* Blackwell Publishing Ltd; 2007 Aug;48(8):444–8.
- 801 97. Concannon P, Whaley S, Lein D, Wissler R. Canine gestation length: variation  
802 related to time of mating and fertile life of sperm. *Am J Vet Res.* 1983  
803 Oct;44(10):1819–21.
- 804 98. Eilts BE, Davidson AP, Hosgood G, Paccamonti DL, Baker DG. Factors affecting  
805 gestation duration in the bitch. *Theriogenology.* 2005 Jul;64(2):242–51.
- 806 99. Evans JM, White K. *The Book of the Bitch.* Interpet; 1997. 1 p.
- 807 100. Gavrilovic BB, Andersson K, Linde Forsberg C. Reproductive patterns in the  
808 domestic dog—A retrospective study of the Drever breed. *Theriogenology.* 2008  
809 Sep;70(5):783–94.
- 810 101. Okkens AC, Hekerman TWM, de Vogel JWA, van Haaften B. Influence of litter size  
811 and breed on variation in length of gestation in the dog. *Veterinary Quarterly.* 1993  
812 Dec;15(4):160–1.
- 813 102. Son CH, Jeong KA, Kim JH, Park IC, Kim SH, Lee CS. Establishment of the prediction  
814 table of parturition day with ultrasonography in small pet dogs. *J Vet Med Sci.* 2001  
815 Jul;63(7):715–21.
- 816 103. Kim BS, Son CH. Time of initial detection of fetal and extra-fetal structures by  
817 ultrasonographic examination in Miniature Schnauzer bitches. *J Vet Sci. The Korean*  
818 *Society of Veterinary Science;* 2007 Sep;8(3):289–93.
- 819 104. Linde Forsberg C, Wikström C, Lundeheim N. Matings by season.
- 820 105. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association  
821 studies. *Nature Genetics.* 2012 Jun 17;44(7):821–4.
- 822 106. Gould SJ. Allometry and size in ontogeny and phylogeny. *Biol Rev Camb Philos Soc.*  
823 1966 Nov;41(4):587–640.
- 824 107. Kleiber M. Body size and metabolism. *Hilgardia.* 1932 Jan;6(January):315–53.
- 825 108. Shingleton A. Allometry: the study of biological scaling. *Nature Education*  
826 *Knowledge.* 2010;3(10):2.
- 827 109. Sacher GA, Staffeldt EF. Relation of gestation time to brain weight for placental  
828 mammals: implications for the theory of vertebrate growth. *Am Nat.* JSTOR;  
829 1974;963:593–615.
- 830 110. Blueweiss L, Fox H, Kudzma V, Nakashima D, Peters R, Sams S. Relationships  
831 between body size and some life history parameters. *Oecologia.* Springer;

- 832 1978;37(2):257–72.
- 833 111. Tuomi J. Mammalian reproductive strategies: a generalized relation of litter size to  
834 body size. *Oecologia*. Springer; 1980;45(1):39–44.
- 835 112. Ross C. The intrinsic rate of natural increase and reproductive effort in primates.  
836 *Journal of Zoology*. Wiley Online Library; 1988;214(2):199–219.
- 837 113. Kihlström JE. Period of Gestation and Body-Weight in Some Placental Mammals.  
838 *Comp Biochem Physiol*. 1972;43(NA3):673–&.
- 839 114. Phillips JB, Abbot P, Rokas A. Is Preterm Birth a Human-Specific Syndrome?  
840 *Evolution, Medicine, and Public Health*. 2015 Jun 14.
- 841 115. Lange K. *Mathematical and Statistical Methods for Genetic Analysis*. New York, NY:  
842 Springer Science & Business Media; 2002. 1 p.
- 843 116. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK:  
844 A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses.  
845 *The American Journal of Human Genetics*. 2007 Sep;81(3):559–75.
- 846 117. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, et al.  
847 Genome sequence, comparative analysis and haplotype structure of the domestic  
848 dog. *Nature*. 2005 Dec 8;438(7069):803–19.
- 849 118. Hoepfner MP, Lundquist A, Pirun M, Meadows JRS, Zamani N, Johnson J, et al. An  
850 Improved Canine Genome and a Comprehensive Catalogue of Coding Genes and  
851 Non-Coding Transcripts. Chadwick BP, editor. *PLoS ONE*. Public Library of Science;  
852 2014 Mar 13;9(3):e91172.
- 853 119. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration  
854 of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*.  
855 2009;4(8):1184–91.
- 856 120. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna,  
857 Austria; 2013.
- 858 121. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and  
859 manhattan plots. *bioRxiv*. Cold Spring Harbor Laboratory; 2014 May 14;:005165.
- 860 122. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-  
861 wide association analysis. *Bioinformatics*. Oxford University Press; 2007 May  
862 15;23(10):1294–6.
- 863 123. Rao TJ, Province MA. A Framework for Interpreting Type I Error Rates from a  
864 Product-Term Model of Interaction Applied to Quantitative Traits. *Genet Epidemiol*.  
865 2016 Feb 1;40(2):144–53.

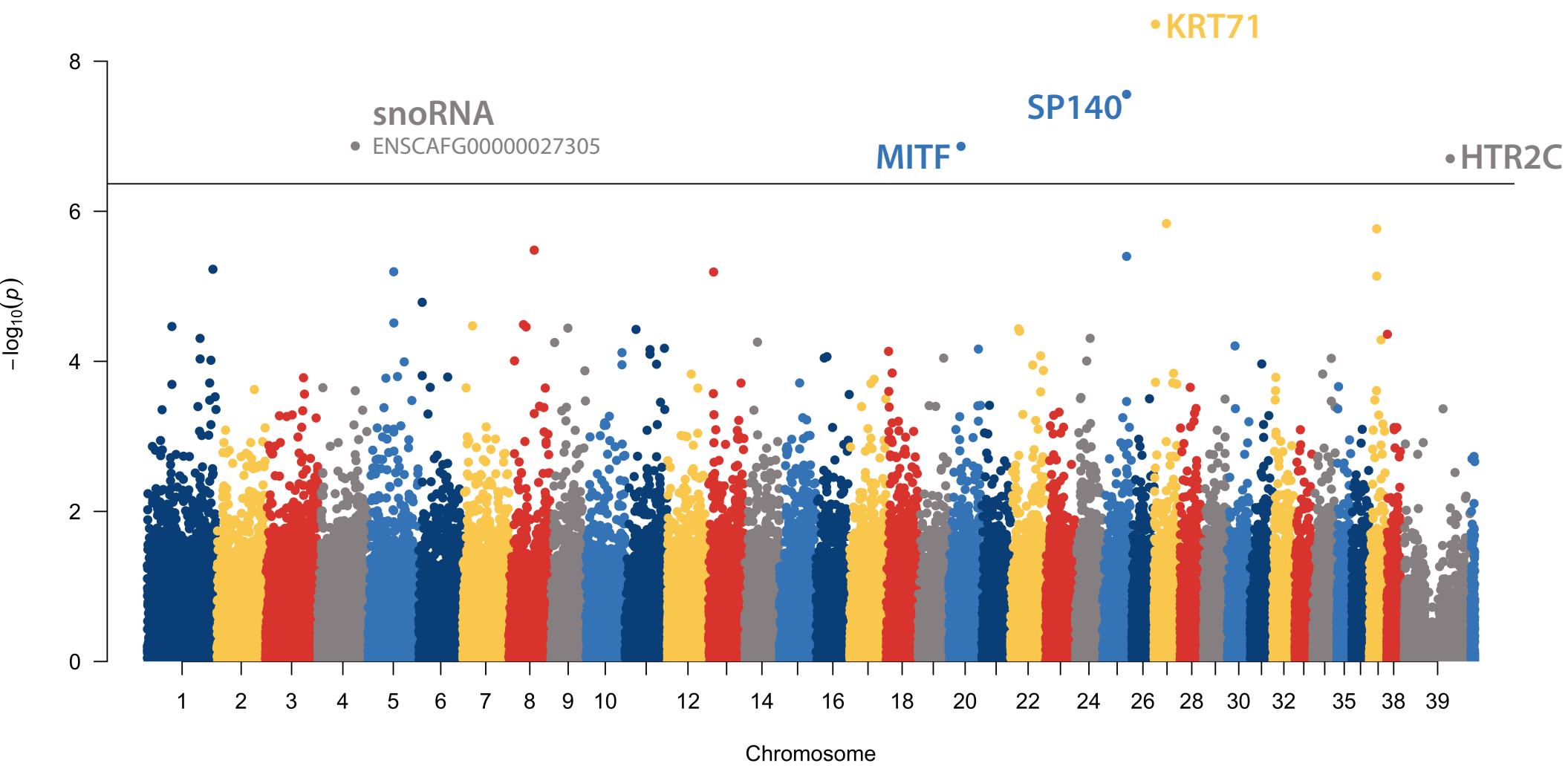
# A. Cesarean Section



# B. Litter Size



# C. Stillbirth



# D. Gestation Length

