Population Genomics of Domestication and Breed Development in Canines in the Context of Cognitive, Social, Behavioral, and Disease Traits

Kristopher J. L. Irizarry and Elton J. R. Vasconcelos

Abstract Dogs are unique because they are known to be the first species domesticated by humans, have the greatest morphological variation among terrestrial mammals, and exhibit unique bonds with humans. Yet, until very recently, the history of domestication and the associated consequences of this artificial selection have been a matter of speculation. Domesticated dogs are the ideal organism to study population genomics of domestication and the impact artificial selection has had on cognitive, social, behavioral, and disease traits. Because dogs have been associated with humans for tens of thousands of years, they are uniquely suited to investigate the genetic basis of selection for dietary adaptation during the agricultural revolution. Through a variety of large-scale genomics approaches, the history and consequences of dog domestication are no longer a matter of speculation. This chapter delves into the ancient origins of human-canine interactions and follows the domestication of wolves into dogs with a particular focus on (a) the selection of phenotypes underlying the strong bond between humans and their companion dogs, (b) the morphological variation underlying dog breeds, and (c) the genetic basis of canine diseases. The historical picture that is beginning to emerge provides a genomics framework for understanding why and how the dog became "our best friend."

K. J. L. Irizarry (✉) · E. J. R. Vasconcelos
The Applied Genomics Center, College of Veterinary Medicine, Western University of Health Sciences, Pomona, CA, USA
e-mail: kirizarry@westernu.edu

1 Introduction

Dogs (*Canis lupus familiaris*) are fondly referred to as "our best friends," and among all organisms on this planet are the species most closely associated with humans. They live in our houses, sleep in our beds, ride in our cars, and even cuddle with us on the couch, while we read and relax. Our shared history extends back tens of thousands of years. This is by far, the longest ever genetic experiment and continues today, with designer dogs selected from crosses of established breeds to produce new and unique combinations of traits. For much of our shared coexistence, the actual impact of domestication and artificial selection has been a matter of speculation.

A wide-held belief is that domestication simply caused dogs to lose their fear of humans. However, the footprints of selection (meaning the specific versions of particular genes selected for in the dog genome during artificial selection) can be detected using population genomics methods. Furthermore, the role these genes play in physiology and biochemistry can be determined using information from prior studies in humans, mice, and other "model" organisms. For example, if one were to identify a gene in wolves (*Canis lupus*) associated with a particular trait and observe that this gene has many different variants in the wolf population, but only a tiny fraction of variants within domesticated dogs, this might provide support for the hypothesis that selection for a particular variant of this gene occurred during artificial selection.

Questions such as "*How are dogs different from wolves?*" and "*What regions of the dog genome encode the traits humans selected for when dogs were domesticated?*" are within the realm of scientific investigation. Population genomics methods provide strategies for decoding the phenotypic consequences of patterns of genetic variation in specific populations. By comparing patterns of genetic variation between two populations (such as wolves and domestic dogs or between Chihuahuas and Great Danes) specific phenotypic differences between the populations can be associated with precise regions of the genome. Such approaches, coupled with comparative genomics and bioinformatics methods, enable us to uncover the particular genes selected during artificial selection and identify the traits these genes encode in the dog genome. Together, this information provides answers to questions about domestication and breed formation.

This chapter presents an informative review detailing genomics aspects of canine domestication and breed formation, with a particular emphasis on cognitive, social, behavioral, and disease traits. The goal of this chapter is to provide a framework for understanding how population genetics and genomics methods have been used to decipher the domestication history and resulting phenotypes that are observed in dogs today. The chapter opens with a brief review of some archeological samples of ancient canids and the results of their genetic analysis. Dogs are the oldest domesticated species and therefore have the longest shared history with humans among all life on the planet. Subsequently, the chapter explores the cognitive and behavioral changes that dogs underwent during their domestication and discusses

a number of studies that have identified the genes underlying these augmented phenotypes that connect them to humans as both companion and working animals.

Next, dog domestication from wolves is presented in the context of sequencing datasets and polymorphic marker analyses. These studies helped elucidate the early history of the dog domestication process. Dogs are known for exhibiting a tremendous amount of phenotypic variation within the species. The chapter then explores the morphological variation and methods employed to deduce the genetic mechanisms underlying this morphological variation. Finally, the chapter delves into clinically relevant phenotypes between specific dog breeds and genes, mutations, and genomic regions underlying these breed-associated diseases. Ultimately, this chapter presents the culmination of our current genetic understanding of canid domestication and provides numerous examples of the specific phenotypes underlying the transformation of ancestral wolves into the dogs we live with today.

2 Time and Place of Dog Domestication

Population genomics methods have offered an unprecedented opportunity to unravel the mysteries underlying dog domestication. These powerful and data-dense genetic approaches have refined our understanding of how dogs transformed from wolves into the hundreds of breeds that exist today. Moreover, through these studies the genomic basis underlying morphological variation between dog breeds is emerging. Through a combination of genetic association studies, whole genome sequencing, and gene expression studies, the veil covering our evolutionary history with dogs has finally been lifted, and the initial discoveries consist of many surprises that, when viewed in the context of "our best friend," make a lot of sense.

2.1 Archeological Evidence

One of the most fundamental and frequently contemplated questions relating humans and dogs is "*When were dogs initially domesticated*?" This is a particularly important question that lies at the heart of the human-animal bond. Ovodov et al. (2011) describe the discovery of 33,000 years old incipient dog remains within the Altai Mountains of Siberia including a complete skull and mandible that were excavated from the site in 1975 (Fig. 1). Evidence of human occupation within the vicinity of the skull and mandible date back approximately 50,000–100,000 years ago and correspond to hunter gatherers that remained in a single location for multiple months at a time (Ovodov et al. 2011).

Fig. 1 33,000-year-old dog skull and mandible represent early stage of canine domestication. (**a**) Aerial view, (**b**) profile, (**c**) palate, (**d**) left mandible, (**e**) left lower tooth row (scale on ruler in cm). Subtriangular hole in the skull is the place of initial sampling for carbon-14 dating in 2007. Originally published in Ovodov et al. (2011)

2.2 Genetic Analysis of Archeological Samples

In 2015 Lee et al. reported the sequencing and phylogenetic analysis of a particular mitochondrial (mt) genomic region, a polymorphic portion of the canine mitochondrial genome that exhibits a 10 bp repeat region that varies by both number of copies and sequence variation between individuals. The data was derived from a 360,000- to 400,000-year-old *Canis* cf. *variabilis* mandible (Fig. 2) obtained from a region in Siberia (Fig. 3) from which multiple ancient and contemporary canid samples have been identified. The study yielded mtDNA region sequence data for all samples investigated leading to the discovery of nine haplotypes. Phylogenetic analysis of the data indicated that the *Canis* cf. *variabilis* sample clustered with other wolf samples from Asia and Russia.

Of particular interest in the Lee et al. (2015) study was the analysis of the haplotypes across ancient wolf samples and contemporary dog breeds. The results indicated that haplotypes obtained from 8,750-year-old samples (site 1 on the map) and 28,000-years-old samples (site 2 on the map) are indistinguishable from

Fig. 2 A 360,000- to 400,000-year-old *Canis* cf. *variabilis* mandible obtained from Siberia. Originally published in Lee et al. (2015)

haplotypes observed in geographically diverse dog breeds that exist today (Fig. 3). A surprising result was that the haplotype observed in the 47,000-year-old canid sample was quite distinct from other wolf haplotypes but differed by only a few mutations from haplotypes observed in the present-day dogs. Taken together, these results provide support for the idea that genetic contributions of ancient Siberian wolves, including possibly *Canis* cf. *variabilis*, may have contributed to the genetic structure of the domestic dog gene pool.

Interestingly, dog domestication appears to have occurred in multiple locations at different times. For example, Thalmann et al. (2013) sequenced the complete mitochondrial genomes of 18 prehistoric canids and compared the results to modern dogs and wolves using maximum likelihood, coalescence, and Bayesian approaches to ascertain phylogenetic relationships. Their findings suggest that contemporary dogs derive their mitochondrial genomes from European canids (Thalmann et al. 2013).

A 2015 study reported by Shannon et al. employed a 185,805-marker genotyping array to investigate the population structure of 4,676 purebred dogs (representing over 160 breeds) and 549 free-ranging village dogs representing 38 countries. The results identified certain geographical subsets of village dogs that appear to be derived almost exclusively from European origins, while village dogs from countries such as Vietnam, India, and Egypt have trace amounts of European admixture, supporting an origin of domestication within Central Asia instead (Shannon et al. 2015).
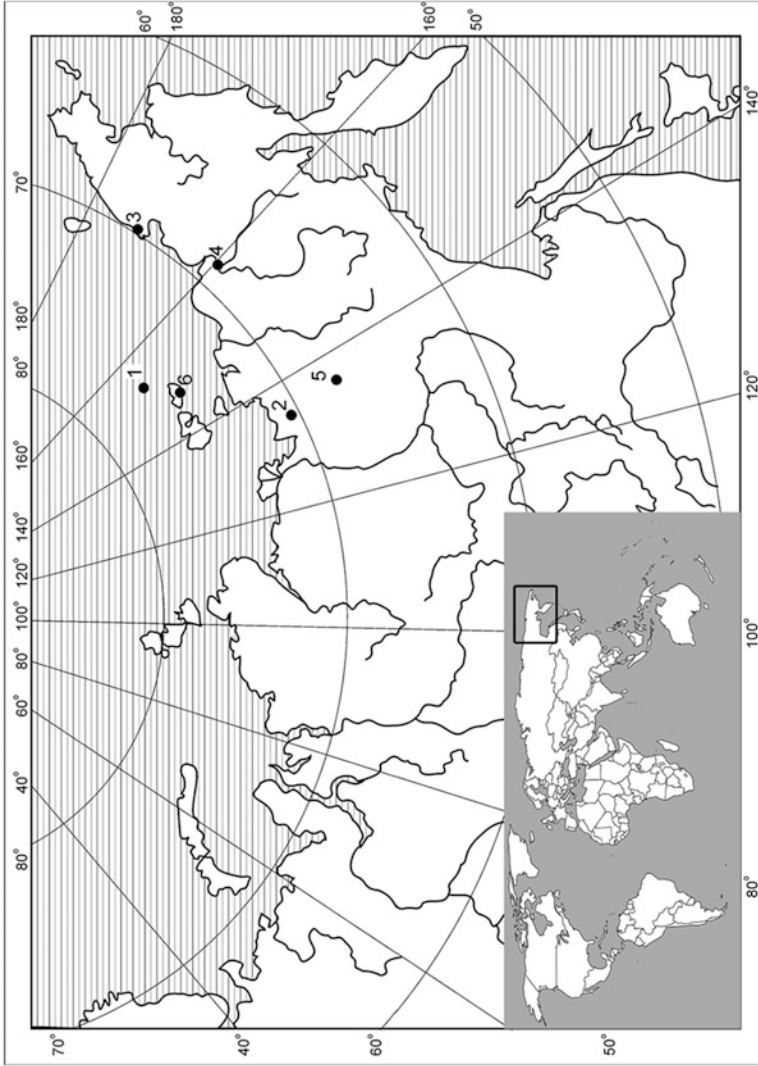
Fig. 3 Region of Siberian Arctic where numerous ancient canid samples have been identified. A number of canid samples have been obtained from six specific sites within this region including: 8,750-year-old *Canis* sp. (site 1 – Zhokhov Island, New Siberian Islands), 28,000-year-old *Canis lupus* (site 2 – Yana RHS, Lower Kolyma River), 1,750-year-old *Canis* sp. (site 3 – Aachim, East Siberian Sea Coast), 47,000-year-old *Canis lupus* (site 4 – DuvanyYar, Lower Kolyma River), 360,000- to 400,000-year-old *Canis* cf. *variabilis* (site 5 – Ulakhan-Suller, Adycha River), and contemporary *Canis lupus* (site 6 – New Siberian Islands). Originally published in Lee et al. (2015)

3 Domestication of Dogs from Wolves

The phenotypic variation among domestic dogs is a consequence of the artificial selection imposed during their domestication and subsequent morphological phenotypic variation that occurred during stratification into different breeds. As of 2018, the American Kennel Club (AKC) recognizes close to 200 distinct dog breeds with additional breeds added each year (http://www.akc.org/). In comparison, the United Kennel Club (UKC) recognizes more than 300 different breeds (https://www.ukcdogs.com) and adds new breeds to the list over time. Similarly, the largest kennel club in the world, Fédération Cynologique Internationale (FCI), currently recognizes close to 350 unique dog breeds (http://www.fci.be). Interestingly, there are dog breeds that are not formally recognized by a breed club. Recently, designer dogs, which are crosses between dogs of different breeds, have gained in popularity. These breeds, lineages, and designer dogs represent pools of dogs that share subsets of genetic variation and together represent one of the most phenotypically diverse species on the planet.

3.1 Early Dog History and Models of Dog Domestication

A recent study of early dog history attempted to characterize the ancestral relationships between dogs and wolves (Lindblad-Toh et al. 2005). This approach used deep genome sequencing of (a) three gray wolves (each centered on a geographical region presumed to correspond to a geographical dog domestication site), (b) two basal dog lineages (the Dingo and Basenji), and (c) the golden jackal (Freedman et al. 2014). In this study the investigators also had access to the Boxer genome because the initial dog genome sequenced published in 2005 was obtained from a female Boxer (Lindblad-Toh et al. 2005). The distribution of samples is illustrated in Fig. 4.

Fig. 4 Geographical distribution of canid samples in genome sequencing study of early dog history. Originally published in Freedman et al. (2014)

Among the data generated across the six canid genomes were 11.2 billion sequencing reads producing over 10.2 million single nucleotide polymorphisms. From this data, the authors inferred effective population sizes based on genome-wide heterozygosity within each genome using a pairwise sequential Markovian coalescent method. By considering an average mutation rate of 1×10^{-8} per generation, the investigators suggest that the dog population underwent a 16-fold reduction over the past 50,000 years. To determine admixture among the genomes, the authors used the nonparametric "ABBA-BABA" test for gene flow between divergent populations. The results of the study were used to construct three models of dog domestication (Fig. 5) that each includes estimates of population divergence and post-divergent gene flow between sample populations.

The three models (Fig. 5) differ in how the ancestral population of wolves ultimately gave rise to different wolf populations and dogs such as the Boxer, the Basenji, and the Dingo. Figure 5a illustrates the model most consistent with the
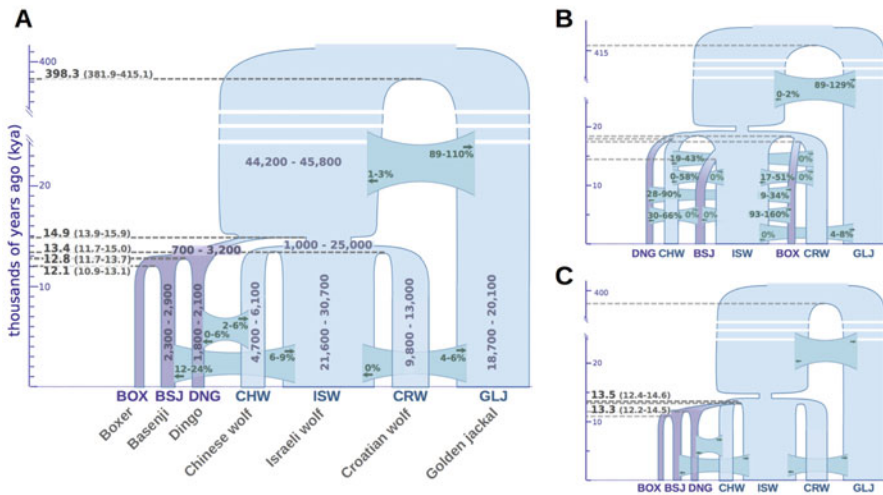

Fig. 5 Models of canine domestication derived from genome sequencing study of early dog history. Divergence times, effective population sizes (N_e), and post-divergence gene flow inferred by *G-PhoCS* in joint analysis of the Boxer reference genome and the sequenced genomes of two basal dog breeds, three wolves, and a golden jackal. The width of each population branch is proportional to inferred population size, and stated ranges of parameter estimates indicate 95% Bayesian credible intervals. Horizontal gray dashed lines indicate timing of lineage divergences, with associated means in bold, and 95% credible intervals in parentheses. Migration bands are shown in green with associated values indicating estimates of total migration rates, which are equal to the probability that a lineage will migrate through the band during the time period when the two populations co-occur. Panels show parameter estimates for (**a**) the population tree best supported by genome-wide sequence divergence, (**b**) a regional domestication model, and (**c**) a single wolf lineage origin model in which dogs diverged most recently from the Israeli wolf lineage (similar star-like divergences are found assuming alternative choices for the single wolf ancestor). Estimated divergence times and effective population sizes are calibrated assuming an average mutation rate of 1×10^{-8} substitutions per generation and an average generation time of 3 years. Originally published in Freedman et al. (2014)

genome-wide sequencing divergence. In this model, the Boxer, Basenji, and Dingo exhibit a lineage from the ancestral wolf population that is distinct from the Chinese wolf, the Israeli wolf, and the Croatian wolf. Figure 5b shows a model in which the Dingo splits off from the Chinese wolf lineage, the Basenji splits from the Israeli wolf lineage, and the Boxer splits from the Croatian wolf lineage. Finally, Fig. 5c represents a model in which all three dogs (Basenji, Dingo, and Boxer) split from the Chinese wolf rather than the ancestral wolf population (as shown in Fig. 5a).

Shearin and Ostrander (2010) provide a quantifiable measure of similarity between dogs and wolves stating that domestic dog differs from its closest ancestor, the gray wolf, by just 0.04% in nuclear protein-coding DNA sequence. In other words, dogs share 99.06% of their protein coding genome with wolves.

4 Evolution and Selection of Cognitive and Behavioral Traits During Canine Domestication

A long-standing question many have asked relating to human domestication of dogs is "*During artificial selection was there any selection for cognitive, behavioral, or communication phenotypes that may have contributed to a strong interspecific bond between humans and their companion dogs*?" The strength of the human-animal bond is so strong that dogs are fondly referred to as humans' "best friends." Subsequently, it seems plausible that artificial selection during domestication may have contributed to divergent phenotypes from wolves that underlie the social interactions between dogs and humans.

4.1 Gene Expression Differences in Brains of Dogs and Wolves

Saetre et al. (2004) investigated the mRNA expression levels of 7,762 genes in dogs, wolves, and coyotes (*Canis latrans*) in three regions of the brain: the hypothalamus, the amygdala, and the frontal lobe. Interestingly, the RNA was obtained from postmortem brains and hybridized to human microarrays. Cross-species micro-array hybridization is inherently challenging, and the extent of sequence divergence between the species (human, dog, wolf, and coyote) contributes to interspecific variation in hybridization efficiency. Nonetheless, the investigators chose to focus on a set of genes that exhibited brain region specificity for one of the three brain regions compared to the other two. Specifically, the selected inclusion criteria required that at least a twofold difference in expression was necessary to consider a gene as brain region specific (Saetre et al. 2004).

In the first set of the gene expression experiments, 156 genes were identified as having region-specific expression in all three species. In a second set of experiments, 114 genes exhibiting expression differences between species within each brain region were identified. Next average interspecies expression differences were determined for all 114 genes. These findings led to the observation that in the amygdala and frontal lobe, average differences in expression were close to 30% and similar across all three species. However, the average expression difference in the hypothalamus was around 20% with a difference between coyotes and wolves of merely 13% (background "noise" was 9%). When wolves and dogs were compared for hypothalamus gene expression, there was an average difference of 24%, and the difference between dogs and coyotes was 22%. Gene Ontology (GO) analysis was performed for the 25 genes that exhibited GO annotation. The results indicated that 25 genes shared annotation of overrepresented GO terms (and only 2 were expected by random chance alone). The enriched terms included neurogenesis, cell-cell signaling, and neurotransmission. Among the genes exhibiting such annotation, many were downregulated in the hypothalamus of dogs. Two of these genes are the neuropeptides *NPY* and *CALCB* implicated in energy regulation, feeding behavior, and the hypothalamic pituitary adrenal (HPA)-associated neuroendocrine stress response. Perhaps domestication of dogs occurred, in part, through genetic variation that modulates gene expression levels in particular regions of the brain underlying stress response phenotypes (Saetre et al. 2004).

4.2 Population Differentiation Between Native Dogs and Wolves

A similar study by Li et al. (2013) employed a pairwise population differentiation between Chinese native dogs and gray wolves. Chinese native dogs are dogs that live as human commensals and were included in the study to capture the genetic structure of dogs prior to the recent stratification associated with breed creation. Furthermore, the authors chose to compare genome-wide divergence between the Chinese native dogs and German Shepherds obtained from Germany. A total of 48,455 SNPs were selected after filtering, and the average distance between the SNPs across the genome was 23 kb. A final set of 1,878 SNPs were identified, corresponding to the top 5% of the distribution, with $F_{ST} > 0.05$ and mean $F_{ST} = 0.63$ between Chinese native dogs and wolves. These SNPs can be considered to be under strong selection (Li et al. 2013).

Gene Ontology biological process enrichment analysis revealed that 347 genes were associated with behavior, and of those, 224 were associated with locomotor behavior. The analysis of SNPs exhibiting highest F_{ST} values between Chinese native dogs and German Shepherds lacked the extent of exacerbated brain expression that was observed among the genes identified between the Chinese native dogs and the wolves. The authors make the case that human artificial selection during the

primary splitting of dogs from wolves was associated with rapid brain evolution. Furthermore, they connect the emergence of dog-specific behaviors during domestication with altered gene expression changes in their brains (Li et al. 2013).

4.3 Whole Genome Sequence Differences Between Dogs and Wolves

Li et al. (2014) compared the published resequenced genomes of three wolves and ten dogs (five ancient dogs, five contemporary dogs) to an additional three wolves and three Chinese native dog genomes that the group sequenced to identify regions of the genome exhibiting the most dramatic differences between dogs and wolves. A common hypothesis associated with dog domestication is that human artificial selection resulted in altered stress response phenotypes, which facilitated dogs and humans living in closer proximity than wolves and humans. Li et al. argue that if stress-response phenotype was "selected" during domestication, one would expect to see evidence of fixed alleles within genes mediating the phenotype to remain fixed today (Li et al. 2014).

Surprisingly, fixed SNPs within the genes *GRIK3*, *GABRA5*, *GRIK2*, *BCL2*, and *MECP2* were identified in the analysis, and GO enrichment identified the following biological processes as the most significantly enriched: adenylate-cyclase-inhibiting G-protein-coupled receptor activity and glutamate receptor signaling pathway. Glutamate is the brain's main excitatory neurotransmitter and regulates behaviors, emotions, cognitive abilities, as well as learning and memory. The gene expression analysis of the GRIK2 gene indicated that it exhibited greater levels of expression in dog prefrontal cortex compared to wolf prefrontal cortex ($p = 0.0006$) (Li et al. 2014).

Although not statistically significant, *BCL2* and *GABRA5* also exhibit changes that distinguish the dog from the wolf. A weighted gene co-expression network analysis revealed that *GRIK2*, *GRIK3*, *GABRA5*, and *MECCP2* exhibit co-expression patterns that place them all in the same coregulatory network. The authors make the case that, during the early stages of domestication, wolves with better learning and memory phenotypes would "come close to human settlements more frequently, acquire greater food resources, and thus had greater opportunities to survive (with little disadvantage). These individuals would exhibit nonaggressive response because they would understand that the presence of humans was harmless, and thus would have a weakened fear response." The authors propose that instead of reduced fear response, domestication of dogs occurred via selection for excitatory synaptic plasticity, which would alter dog behavior and cognition to the point where dogs could learn the meaning of human gestures and respond more favorably to human commands (Li et al. 2014). The idea that artificial selection during domestication altered the canine brain to enhance dog memory is an exciting and potentially transformative event in human-animal history.

5   Genetic Effects of Dog Domestication

Domestication events can create bottlenecks and consequently reduce genetic diversity, reduce effective population size, and increase inbreeding. Understanding the relationship between the genomic signals observed in the data and the evolutionary mechanisms that contributed to those signals is critical if one hopes to understand how the domestication and selective breeding history of contemporary dog breeds exploited the morphological plasticity encoded in the ancestral canine genome. Boyko et al. (2010) suggest long runs of homozygosity (ROHs) are the result of inbreeding associated with recent selection events, such as breed formation. In contrast, the authors attribute haplotype diversity and linkage disequilibrium (LD) occurring across genomic scales less than a megabase as indicative of more ancestral population processes (Boyko et al. 2010).

The first question addressed by Boyko et al. (2010) was to investigate genomic signatures of canine demographic history by analyzing (1) the pairwise SNP LD, (2) the haplotype diversity across 15-SNP windows, and (3) the extent of ROHs greater than a megabase. They discovered that although the LD extends over 1 Mb within every breed assessed across the entire population of dogs, it decays very quickly. This observation implies that identity-by-descent (IBD) segments are shared across numerous breeds and are quite small. The ROHs observed were longer and occurred more frequently in breed dogs than wolves or the village dogs. Individuals from almost all breeds exhibited between 10 and 50 ROHs greater than 10 Mb. The exception was the Jack Russell Terriers, which showed fewer ROHs and higher levels of genetic diversity than the other breeds (Boyko et al. 2010).

Autozygosity, which occurs when both chromosomes in a diploid organism are derived from the same ancestor, indicates that inbreeding has occurred. Current models suggest "inbreeding depression" is an increase in autozygosity coupled with an increased risk in homozygosity at rare, partially recessive, deleterious mutations. To investigate the impact of autozygosity, it is important to accurately identify real autozygous ROHs from the larger set of often non-autozygous ROHs in a sample (Boyko et al. 2010). Non-autozygous ROHs, stretches of homozygous SNPs that are actually heterozygous at unmeasured variants, are less likely to contain rare, partially recessive, deleterious mutations in homozygous form. Subsequently, an important criterion for defining ROHs – rather than SNP-by-SNP homozygosity – is to assess autozygosity. It is important to identify ROHs that are not autozygous and are identical-by-state (IBS) from ROHs that are autozygous and are identical-by-descent (IBD) (Howrigan et al. 2011).

According to Boyko et al. (2010), autozygosity was detected at high levels in all breeds with Jack Russell Terriers having the lowest average autozygosity (7.5%) and Boxers having the highest (51%). Interestingly, only a few breeds contained genomic regions that were autozygous in all breed members genotyped at the megabase scale. The exception was Basenjis, which showed evidence of high haplotype diversity coupled with high autozygosity. Together these two conditions are suggestive of a recent genetic bottleneck following breed formation that caused

greater levels of inbreeding than would otherwise be expected in the population. According to the breed history, Basenjis in the United States were derived from a relatively small founder population. Linkage disequilibrium (LD), associated with regions of chromosomes encoding shared alleles from common ancestors along a chromosome, is known to extend greater genomic distances within breeds than it does among breeds or within wolves. The analysis performed by the authors indicates that the between-breed LD is much greater than wolf LD which provides support for a bottleneck in dogs during domestication (Boyko et al. 2010).

These results support the idea that dramatic genomic selection occurred within the dog genome on multiple time scales. One time scale, for example, corresponds to an ancient domestication selection process when dogs were selected for affiliation with humans. Afterwards, a more recent breed-radiation selection process occurred where closed breeding pools were created to transform the ancestral genetic variation into breed-specific pockets of genetic and morphological phenotypic uniformity.

6 How Did Domestication-Modulated Oxytocin Mediated Phenotypes

6.1 Oxytocin-Mediated Social Phenotypes in Dogs

The neuropeptide hormone, oxytocin, has a well-established role underlying social bonding in mammals where, through evolution, it has mediated hierarchical social relationships as well as organization of social interactions. In humans, oxytocin coordinates parental responses after physical contact with offspring, interactions between sexual partners, interactions with friends, and empathetic interactions with strangers (Feldman 2017).

Romero et al. (2014) described a prosocial role for oxytocin in dogs. They suggested that oxytocin facilitates prosocial interactions among dogs and humans. Furthermore, they make the point that evolutionary selection pressure may have contributed to the maintenance of neurological mechanisms associated with social bonding due to the adaptive value of long-lasting social relations (Romero et al. 2014).

6.2 Genetic Variation in Dog Oxytocin Receptor

The role of oxytocin signaling in the human-animal bond suggests that it is possible that domesticated dogs were artificially selected for more affiliative relationships with humans through allelic variation within genes mediating oxytocin signaling. And indeed there is evidence for considerable genetic variation within the oxytocin receptor in canids, as well among different dog breeds (Kis et al. 2014; Bence et al. 2017).

Kis et al. (2014) investigated three polymorphisms within the receptor located within either the 5′ UTR or the 3′ UTR of the gene. They genotyped 103 Border Collies (46 males, 57 females), consisting of two subpopulations (59 from Hungary and 44 from Belgium). Additionally, they genotyped a single population of 104 German Shepherd dogs (58 male, 46 female) and assessed behavioral phenotypes across five specific tests: (1) greeting the dog, (2) separation from owner, (3) problem-solving, (4) threatening stranger, and (5) owner hiding from dog. The study results demonstrated evidence of an association between the G-allele of -212AG polymorphism and the behavioral phenotype of decreased owner proximity seeking in both breeds. Additionally, the authors report an association of the rs8679684 polymorphism with friendliness; however, the breeds exhibited divergent associations with the *A*-allele in German Shepherds exhibiting higher friendliness phenotype scores, while in Border Collies, the *A*-allele was linked to decreased friendliness (Kis et al. 2014). Note that the -212AG polymorphism was subsequently renamed to the -213 AG polymorphism as the genomic coordinates for the canine oxytocin receptor were refined.

Bence et al. (2017) characterized nine oxytocin receptor polymorphisms in four different canid species. Their study included three novel oxytocin receptor polymorphisms identified through direct sequencing of the gene and regulatory regions in two Eurasian gray wolves, four North American timber wolves, three Beagles, three Border Collies, three German Shepherds, three Golden Retrievers, and three Siberian Huskies. This sequencing led to the identification of -74C/G, 18575C/T, and a microsatellite marker occurring between positions 18772–18792. They also included the three polymorphisms reported in 2014 by Kis et al., -213A/G, 19208A/G (previously called -212A/G and 19131A/G, respectively), and rs8679684. Additional three polymorphisms were identified in public database searches (Bence et al. 2017). Allele frequencies were assessed in 689 purebred dogs (70 Beagles, 144 Border Collies, 128 German Shepherds, 43 Golden Retrievers, 22 Groenendaels, 32 Hungarian Vizslas, 49 Labrador Retrievers, 40 Malinois dogs, 138 Siberian Huskies, and 23 Tervurens) as well as 42 wolves (34 Eurasian gray, 6 North American timber, 2 Alaskan), 6 golden jackals, 8 Dingos, and 45 Asian street dogs.

The results revealed that only the -213A/G G-allele, -94C/T C-allele, -74C/G C-allele, -50C/G C-allele, rs22927829 T-allele, rs8679684 T-allele, and 19208A/G G-allele were detected in all four species. Interestingly, -213A/G A -allele, -50C/G G-allele, and 19208A/G A-allele are only found in wolf and dog, with the wolf having a higher allele frequency than the dog in each case. The rs22927829 A-allele was only detected in dog and Dingo, while the rs8679684 A-allele was found only in dogs. Across the dog breeds and wolf, Bence et al. (2017) reported that only two of the polymorphisms exhibited evidence of both alleles in Border Collie, Golden Retriever, Labrador Retriever, Hungarian Vizsla, Beagle, Tervuren, Groenendael, Malinois, German Shepherd, Husky, and wolf (-94T/C, -74C/G). The -213A/G polymorphism, for which the G-allele was implicated in owner proximity seeking (Kis et al. 2014), lacked evidence of the G-allele in Tervuren and Groenendael breeds. These results underscore the notion that phenotypic variation in social behavior may exist across dog breeds (Bence et al. 2017).

6.3 Visual Communication and Oxytocin

Nagasawa et al. (2015) investigated the physiological consequence of gazing behavior between dog and owner. The rationale for this study was based on the idea that human-like modes of communicating, such as mutual gaze, may have been selected in dogs during domestication by humans. The authors refer to maternal oxytocin levels rising in human mothers when mother-infant gazing occurs. They designed experiments to test the hypothesis that an oxytocin positive feedback loop may be induced by gaze between dogs and their human owners. From the results of their experiments, the authors suggest the existence of a self-perpetuating positive feedback loop mediated by oxytocin in the human-dog bond. The authors characterize the human-dog bond as being similar to the maternal-infant bond because both bonds are associated with oxytocin positive feedback loops across the bond members. Nagasawa et al. (2015) extrapolate from their results and suggest that gazing behavior between dog and owner over thousands of years of domestication and cohabitation conferred social rewarding effects to both humans and dogs. They further point out that this oxytocin release, in both the dog and the human, would result in a deepening of the mutual relationship and further promote interspecies bonding (Nagasawa et al. 2015). They also examined whether an oxytocin loop may have been acquired during dog domestication or whether it is shared among canids that did not undergo domestication by employing hand-raised wolves in their research (Nagasawa et al. 2015). The wolves did not exhibit long periods of gazing at humans. The authors interpret this finding to mean that wolves do not engage in mutual gaze as a means of social communication and interaction with humans. Furthermore, the authors point out that in wolves, eye contact is considered a threat among conspecifics and wolves generally avoid eye contact with humans (Nagasawa et al. 2015).

6.4 Interbreed Differences in Oxytocin-Mediated Phenotypes

Dog breeds differ in social behavior in response to oxytocin. Kovacs et al. (2016) demonstrated the existence of interbreed differences in social behavior associated with intranasal oxytocin in two dog breeds (Siberian Husky and Border Collie) representing distinct genetic lineages. Kovacs et al. genotyped the dogs (18 Siberian Huskies and 16 Border Collies) on the -213A/G oxytocin receptor polymorphism and identified an association between the dog's genotype and social behavior (Kovacs et al. 2016).

The path to canine domestication resulted in selection for traits contributing to enhanced social bonding with humans and increased perception of human communication and nonverbal gestures. The acquisition of these traits allowed dogs to inhabit a unique social niche among humans. Persson et al. (2016) employed a high-density SNP chip and identified *SEZ6L* as a gene exhibiting an association with

variation in social traits. *SEZ6L* has been implicated in autism, a phenotype in which social interaction and communication deficits occur. Other genes located in proximity to the identified haplotype block in the study include *ARVCF*, which has been linked to schizophrenia, and *TXNRD2* and *COMT*, two genes that play roles in schizophrenia and social disorders (Persson et al. 2016).

7 Regions of the Dog Genome Exhibiting Evidence of Positive Selection During Domestication

7.1 Identification of Positively Selected Genes in Dogs Compared to Wolves

Wang et al. (2013) performed a genomic analysis to identify genes that exhibit evidence of positive selection. They highlight the point that artificial selection acting on dogs occurred in two phases. The first phase was defined by the domestication of dogs from wild canids. These descendants of wolves shared living environments with humans and subsequently shared human dietary resources. The second phase was much more recent, occurring over the last few hundred years when morphological variation was created leading to the diverse array of breeds and the physical phenotypes that define them. Wang et al. (2013) suggest that genes selected during the first phase should be shared among all dogs today and designed the experimental approach in this context. Specifically the authors looked for regions of the genome that contain relatively low levels of diversity between dogs and high levels of diversity between wolves and dogs. Regions of the dog genome that contained low levels of diversity in wolves were excluded from the analysis to prevent the identification of genomic regions exhibiting low diversity in dogs that were inherited directly from wolves without selection during domestication (Wang et al. 2013).

Among a set of 17,661 orthologous gene pairs between dogs and humans, 1,708 and 233 genes exhibited evidence of positive selection for humans and dogs, respectively. Gene Ontology enrichment analysis identified terms such as "regulation of digestion," "negative regulation of intestinal phytosterol absorption," "regulation of lipid transport," "axon," "neuron projection," "cell projection," "gamete generation," "sexual reproduction," and "reproductive process in a multicellular organism." These terms are particularly interesting because they reflect three major themes of evolutionary selection during the initial phase of dog domestication: (1) digestion, (2) reproduction, and (3) neurological process (Wang et al. 2013).

Strikingly, among these three functional categories, the authors identified orthologous genes between dogs and humans that show evidence of positive selection in both species. Those genes include *ABCG5*, *ABCG8*, *PLA2G10*, and *PRSS1* associated with nutrition. The genes *GRM8* and *SLC6A4* were identified within the neurological process group. Among the genes implicated in reproduction, *BFAR*, *BRE*, *ITGB1*, *MET, STK17B*, and *ZMYM2* were identified as being positively

selected in both dogs and humans and are involved in cancer, apoptosis, and cell cycle as genes within the reproductive category. Of particular interest are the neurological genes, *GRIM8* and *SLC6A4*, which correspond to glutamate receptor metabotropic 8 and the serotonin transporter, respectively. These genes modulate phenotypes associated with autism and personality traits in humans (Wang et al. 2013).

7.2 Selection for Enhanced Starch Digestion in Domestic Dogs

A major event in the domestication of dogs was the selection for a starch diet. Axelsson et al. (2013) used approximately four million SNPs to identify multiple genes associated with starch digestion and fat metabolism that exhibit evidence of selection in dogs (Axelsson et al. 2013). Specifically, the authors identified ten genes implicated in digestion and fat metabolism that were associated with specific mutations found in dogs. These results provide genetic evidence that domesticated dogs adapted to survive on starch-rich diets compared to the carnivorous diets of their wolf ancestors.

In a follow-up study, Arendt et al. (2014) found that high amylase activity in dogs was correlated with pancreatic amylase (*AMY2B*) copy numbers in the genome. The authors characterized the distribution of *AMY2B* copy numbers across 20 breeds and showed that considerable heterogeneity in *AMY2B* copy number exists across dog breeds, ranging from 6 to 14 copies per genome. Dogs living with humans that were exposed to agricultural advances during the prehistoric rise of agriculture benefitted from these dietary resources. Arendt et al. (2016) determined that adaptation to starch diets did not occur early in dog domestication but rather occurred in subpopulations that were exposed to starch-rich diets. Their results show high levels of *AMY2B* copy numbers in most domesticated dogs but relatively few in dogs originating from the Arctic. This is consistent with the historical geographic spread of agriculture (Arendt et al. 2016).

Reiter et al. (2016) demonstrated that positive selection continued to act on dogs that were exposed to starch-rich diets well after dog domestication had occurred. The authors analyzed the relationship between starch-rich diets and dog breeds to gain a better understanding of the relationship the dietary starch played in *AMY2B* copy numbers. Their results demonstrate that dogs exposed to dietary starch exhibit higher allele frequencies of diploid *AMY2B* repeats. This relationship can be seen within specific dog breeds, such as the Shar Pei and Pekingese (exposed to high-starch diets) compared to the Siberian Husky and Alaskan Malamute (exposed to low-starch diets) as illustrated in Fig. 6 (Reiter et al. 2016).
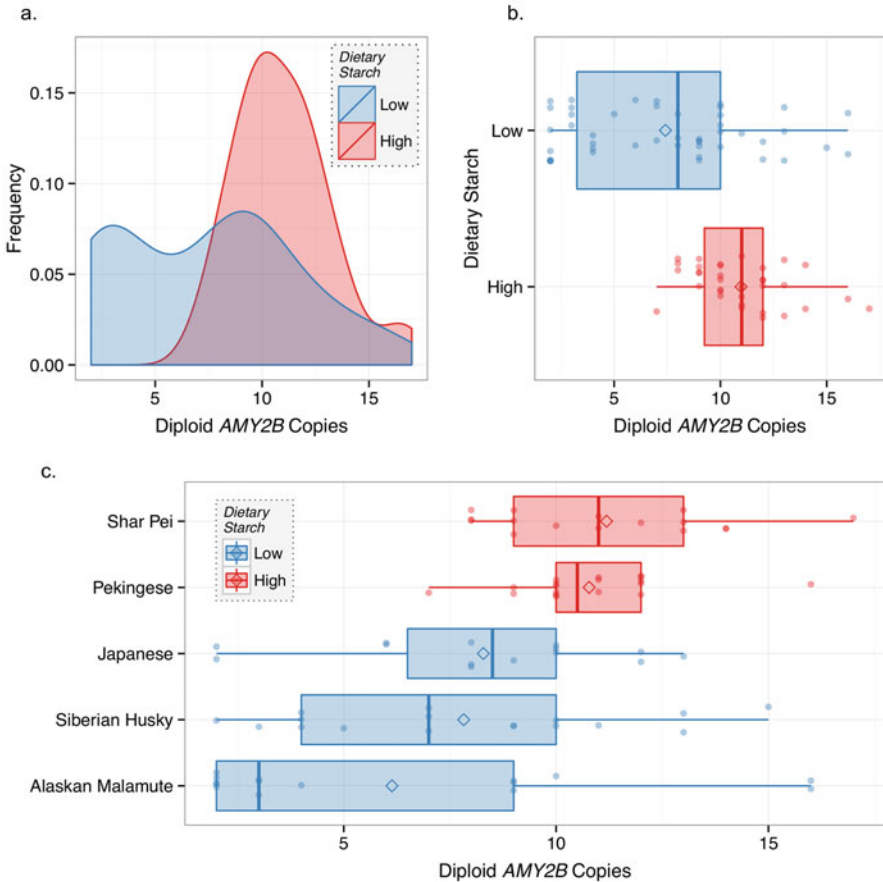
Fig. 6 Diet and AMY2B copy number variation. (**a**) Density plot of ddPCR diploid AMY2B copy number for dogs that traditionally consumed high-starch diets and low-starch diets. Density reflects frequency with which a given diploid copy number appears in each population. (**b**) Tukey boxplot of diploid AMY2B copy number for dogs that traditionally consumed high-starch diets and low-starch diets. (**c**) Tukey boxplot of diploid AMY2B copy number for specific dog breeds that traditionally consumed high-starch diets and low-starch diets. Originally published in Reiter et al. (2016)

7.3 Functional Polymorphisms Exhibiting Fixed Alternative Alleles in Dogs and Wolves

Another study investigating genomic regions targeted by selection during dog domestication was described by Cagan and Blass (2016). Their approach leveraged searching a comprehensive canine polymorphism database to identify polymorphic markers that are highly differentiated between wolves and dogs. Their approach led to the identification of 11 genes for which functional variants are fixed for alternative alleles in dogs and wolves. A pathway analysis of the genomic regions

containing the polymorphic markers with $F_{ST} > 0.75$ identified "adrenaline and noradrenaline biosynthesis pathway," "axon guidance mediated by netrin," "dopamine receptor-mediated pathway," "nicotine pharmacodynamics pathway," "alpha adrenergic pathway," and "gonadotropin-releasing hormone receptor pathway." The authors point out that each pathway was represented by multiple genes. Furthermore, computational analysis suggested that within each of the pathways, there are genes with putatively functional variants (Cagan and Blass 2016).

The authors state that domestication of dogs likely selected for reduced fight or flight responses, which are, in part, mediated by pathways such as "adrenaline and noradrenaline biosynthesis pathway" (nine genes with potentially functional variants), "dopamine receptor-mediated signaling pathway" (eight potential functional variant genes), "alpha adrenergic receptor signaling pathway" (five potential functional variant genes). The identification of neuro-related pathways further lends support to the idea that behavioral phenotypes were selected during the initial phase of dog domestication when wolves and dogs first began diverging (Cagan and Blass 2016).

8 Genetic Structure of Dog Breeds

After dog domestication, the next most frequently pondered questions about dogs are: "*How were the different breeds created?*" and "*What components of the genome are responsible for the morphological phenotypes that define these breeds?*" Answers to these questions lie at the heart of many population genetics/genomics studies carried out on dogs.

8.1 Dog Genome Sequence and Genetic Diversity

Studies revealing the sequence of the dog genome and canine genetic variation have provided considerable information about the population structure of purebred dogs and the relationship between different breeds. The dog genome sequence, derived from a female Boxer, was published in 2005 (Lindblad-Toh et al. 2005). The Boxer was selected due to the decreased heterozygosity within the breed and an expected easier genome assembly process than would be expected for a dog with much greater heterozygosity (Lindblad-Toh et al. 2005).

The genome was sequenced with the whole genome shotgun approach resulting in over 31 million sequence reads corresponding to $7.5\times$ coverage of the ~2.4 billion base pair genome. The assembly was anchored to dog chromosomes with data derived from previously constructed cytogenetic and radiation hybrid maps. The resulting genome sequence enabled the identification of an initial set of 19,300 protein-coding genes. An analysis of 13,816 1:1:1 orthologs between human, mouse, and dog provided lineage-specific data on synonymous (K_S) and

non-synonymous (K_A) changes. This allowed the investigators to calculate the K_A/K_S ratio, which provides a measure of the strength of selection acting on protein coding genes. As part of their analysis, the authors determined the median K_A/K_S ratios and discovered that the ratio differed substantially across each of the lineages. Their results placed the K_A/K_S ratio within the dog lineage in between the mouse and human lineages. The authors relate this finding to the population genetics theory that associates strength of purifying selection with increased effective population size. Their results are consistent with this theory as smaller mammals (such as mouse) tend to have larger effective population sizes (Lindblad-Toh et al. 2005).

To better understand canine genetic diversity, three distinct SNP datasets were analyzed. Lindblad-Toh et al. (2005) identified a total of 770,000 SNPs within the Boxer genome. The authors also compared a previously assembled 1.5× coverage draft sequence of the poodle genome (Kirkness et al. 2003) to their sequence of the Boxer. The comparison identified 1,460,000 SNPs between the two dog breeds (Kirkness et al. 2003). Additionally, Lindblad-Toh et al. (2005) generated shotgun sequencing data from 9 diverse dog breeds, 4 gray wolves, and 1 coyote using 22,000 sequencing reads from each that resulted in a set of 440,000 SNPs. A 1,283 subset of these SNPs were validated by resequencing which indicated a true positive rate of 96% (Lindblad-Toh et al. 2005).

8.2 Single Nucleotide Polymorphisms in the Dog Genome and Inference of Bottleneck Events

A comprehensive SNP map was constructed from the above three SNP datasets resulting in a final SNP map of more than 2.5 million SNPs. On average, any two dogs will have a single nucleotide polymorphism within approximately every thousand base pairs between members of different breeds, while members of the same breed will have a SNP within 1,500 bp of their genomes. According to their analysis, the gray wolf (1/580 bp) and the coyote (1/420 bp) exhibit greater genetic variation than the Boxer. Within the Boxer assembly itself, a SNP occurs within roughly every 3,000 bp. Based on their identification and analysis of SNPs, the authors conclude that a set of 10,000 SNPs is sufficient for genetic association studies in dogs (Lindblad-Toh et al. 2005).

As part of their analysis, Lindblad-Toh et al. (2005) modeled the population history of the domestic dog. Specifically, they built a mathematical model in which a dog population experienced both an ancient and a recent bottleneck. The results of their coalescent method fit well with their genetic data when they set the ancient bottleneck to 9,000 generations ago (27,000 years ago), with a population size of 13,000 and an inbreeding coefficient of $F = 0.12$, and to the more recent breed-creation bottleneck 30–90 generations ago (90–270 years ago). The authors also used the modeling approach to generate estimates of breed-specific bottlenecks that were consistent with known histories of the breeds. The breed that exhibited the

poorest fit to the two bottleneck model was the Akita which was created in Japan about 450 generations ago and then underwent a subsequent bottleneck in the 1940s when it was introduced into the United States (Lindblad-Toh et al. 2005).

8.3 Number of Dog Breeds

Worldwide, there are over 400 recognized dog breeds. The American Kennel Club recognizes 192 dog breeds as of 2018, and each year one million dogs are registered by the AKC with over half of all annual AKC dog registrations corresponding to just 10 breeds.

8.4 Microsatellite Analysis of the Genetic Structure of 85 Dog Breeds

An initial analysis of 85 dog breeds (genotyping 5 unrelated dogs from each breed) was conducted using 96 microsatellite markers that spanned the canine genome with an average density of approximately 30 Mb (Parker and Ostrander 2005). The results indicated that a purebred dog could be assigned to its breed of origin 99% of the time. During the analysis, it was discovered that almost 40% of all genetic variation occurring in dogs is detectable when comparing dogs across breeds, for example, when comparing a Great Dane to a Chihuahua versus comparing one Chihuahua to another Chihuahua. This is considerably greater than what has been observed in humans, where just 5–10% of all human genetic variation occurs between populations and races. The genotyping data was used to cluster the 85 breeds based on genetic similarity. Although most breeds mapped cleanly to a single cluster, some breeds such as Australian Shepherd, Bichon Frise, Flat-Coated Retriever, Great Dane, Lhasa Apso, and Pug mapped to more than one cluster (Fig. 7) (Parker and Ostrander 2005).

8.5 Genetic Diversity Differences in Dog Breeds

Quignon et al. (2007) assessed the extent of genetic diversity inherent in Bernese Mountain Dogs (BMD), Flat-Coated Retrievers (FCR), Golden Retrievers (GR), and Rottweilers (ROT) sampled in equal proportions from the United States and Europe. The goal of the study was to better understand how genetic variation within dogs of the same breed varies by geographic location. Genetic studies in dogs can be confounded by population stratification resulting in false-positive associations when population substructure exists within a breed. This can be particularly problematic when studies are designed assuming that all dogs within a breed share the same level
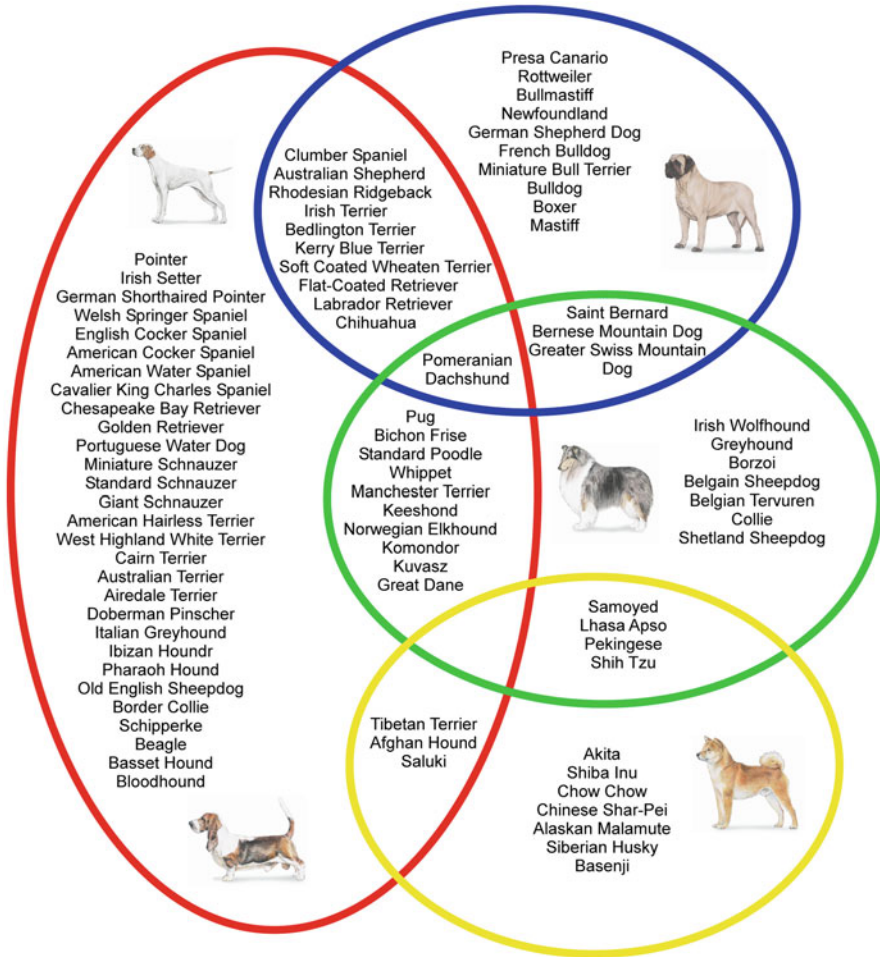
Fig. 7 The population structure of 85 dog breeds. The dataset includes five unrelated dogs from each of the 85 breeds that have been genotyped using 96 (CA)n repeat-based microsatellites that spanned the dog genome at an average density of 30 Mb. Clusters were obtained using the computer program Structure, which implements a Bayesian model-based clustering algorithm that attempts to identify genetically distinct subpopulations based on patterns of allele frequencies. Four distinct clusters described by Parker et al. are depicted as colored circles: cluster one is yellow, cluster two is blue, cluster three is green, and cluster four is red. Breeds associated with each cluster are listed within the appropriate circle, and examples of breeds are shown in the pictures. Some breeds show similarity to more than one cluster and are listed in the overlapping space. Originally published in Parker and Ostrander (2005)

and type of genetic variation. A set of 722 SNPs from four loci on chromosome 1 was genotyped in 120 dogs (Quignon et al. 2007). The investigators determined that the GR exhibited the greatest number of polymorphic SNPs (66.6%), while the fewest polymorphic SNPs were detected in the BMD. The FCR had 57.7% polymorphic SNPs, and the ROT had 54.4% polymorphic SNPs (Quignon et al. 2007).

The finding that dog breeds are not homogenous populations underscores the importance of population substructure when considering case-control genetic association studies. The authors state that variation in allele frequencies can arise through a population's genetic history, ancestral geographical distributions, mating practices, and both reproductive expansions and bottlenecks. Moreover, Quignon et al. indicate that besides population stratification arising from variation in geographical origin, artificial selection during breeding for phenotypic traits such as coat color, herding, hunting, olfactory capabilities, memory, and cognitive ability can also result in undetected population structure when those breeds are used in genetic studies (Quignon et al. 2007). This study highlights the fact that although members of a dog breed may share similar physical traits, each dog is genetically a unique individual.

8.6 Genome-Wide Genetic Structure and Evolution of Dogs Versus Wolves

Vonholdt et al. (2010) carried out a genome-wide analysis of 48,000 SNPs in 912 dogs (representing 85 breeds) and 225 gray wolves (across 11 globally distributed populations). The goal of the study was to gain a better understanding of the evolutionary and geographical history that gave rise to the dramatic diversification of phenotypes observed in dogs today. The authors used Bayesian clustering methods to identify any dog breeds that may have evidence of admixture with wolves. A relatively small set of breeds, considered ancient dog breeds, were identified and include breeds such as Afghan Hound, Akita, Alaskan Malamute, Basenji, Chinese Shar Pei, Chow Chow, Dingo, and Siberian Husky to name a few. Based on historical information, these ancient dog breeds have origins dating back more than 500 years ago (Vonholdt et al. 2010).

To determine the main contribution of genetic diversity in domestic dogs, Vonholdt et al. (2010) considered whether a single wolf population clustered with dogs in neighbor-joining trees by taking into account allele sharing of individual SNPs, 5-SNP haplotypes, and longer multi-SNP haplotypes for individuals and breed groupings. Their results indicated that only for individual SNPs and 5-SNP haplotypes Middle and Near Eastern gray wolves clustered with dogs. Moreover, in this analysis all other wolves clustered together as a single genetic entity separate from dogs. Then they tested whether haplotypes sharing of modern and ancient dog breeds could be associated with any distinct wolf populations. For this analysis, North American wolves were used as a negative control based on existing models of dog domestication excluding North America as the center of dog domestication (Vonholdt et al. 2010).

The results demonstrated that the extent of shared haplotypes between dogs and North American wolves was lower than sharing between dogs and Old-World wolves. More importantly, they discovered that for 5-SNP haplotypes, sharing was greater between Middle Eastern wolves and modern dog breeds than between other

populations of wolves. For longer multi-SNP haplotypes, the authors report that most breeds exhibit the greatest haplotype sharing with Middle Eastern wolves, including geographically diverse breeds such as Basenji, Bassett Hound, Borzoi, and Chihuahua. In separate analysis, the Akita, Chinese Shar Pei, Chow Chow, and Dingo shared most strongly with Chinese wolves. Finally, the authors note that for the 5-SNP and longer multi-SNP haplotype analyses, the Basenji shared the most haplotypes with Middle Eastern gray wolves than any other domestic dogs. It is worthwhile to point out that Basenjis are a dog breed having a Middle Eastern origin. The authors interpret this result as indicating that the Basenji had a large effective population size early in domestication or, alternatively, they have been recently backcrossed with wolves. Taken together, the authors conclude that the Middle East is the main source of genetic diversity in dogs with possible minor contributions derived from Europe and Asia (Vonholdt et al. 2010).

Vaysse et al. (2011) described a comprehensive high-density genotyping analysis of genomic regions exhibiting evidence of selection in 509 dogs across 46 diverse breeds and 15 wolves using 170,000 evenly spaced SNPs. Evolutionary relationships between the sampled subjects were assessed by building a neighbor-joining tree from the genetic distances in the comprehensive genotyped dataset (Fig. 8). Visualizing this tree led to the following conclusions: (1) dogs from the same breed clustered together as is expected from closed gene pool breeding groups, (2) relatively no structure is present within the breeds which is consistent with modern dog breeds arising from a common set of ancestors rather quickly, and (3) the internal branches for Boxer and wolf are longer than those for other breeds which make sense because SNP discovery occurred using genomic sequence data from the Boxer genome and the longer wolf branches likely represent greater evolutionary distance compared to the other dog breeds (Vaysse et al. 2011).

9 Genomic Basis for Morphological Variation Between Dog Breeds

Although dog domestication began at least 15,000 years ago, it wasn't until the Victorian era, roughly 200 years ago, that artificial selection for breed standards in dogs first began. The phenotypes observed in the breeds of today represent extremes of morphological variation (Fig. 9) (Shearin and Ostrander 2010).

Phenotypic variation across breeds is the consequence of a variety of physical traits associated with numerous anatomical regions. Variation in skeletal morphology is associated with differences in body size, leg size, and skull shape between breeds. Tremendous variation in hair phenotypes gives rise to differences in coat texture, length, and color within different breeds (Fig. 9).
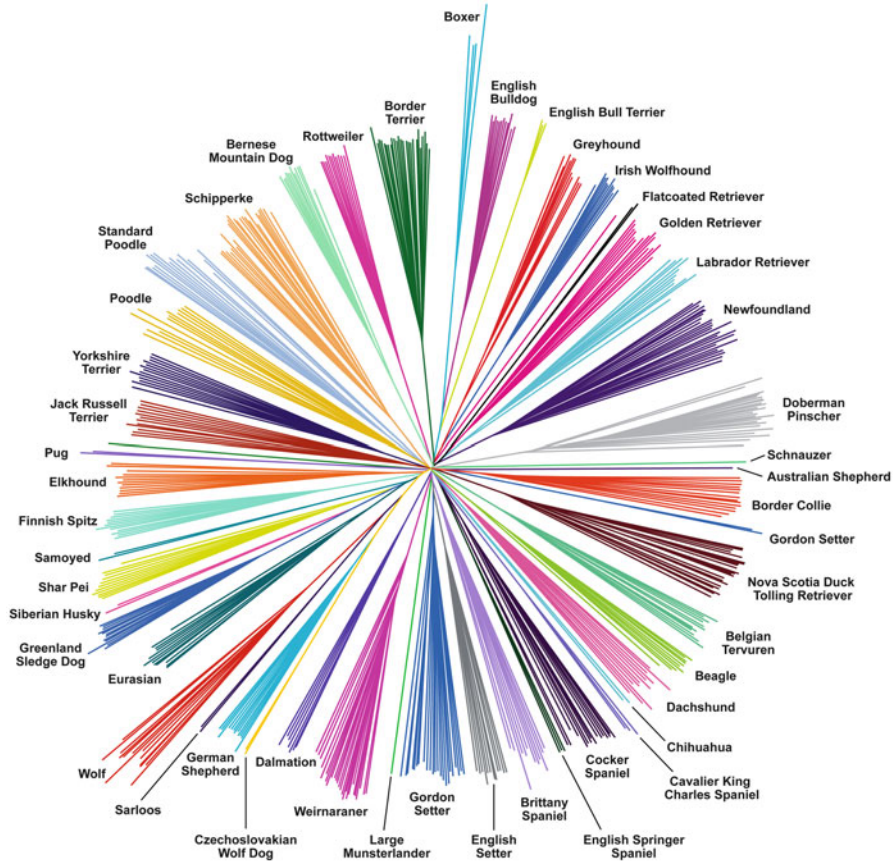
Fig. 8 Neighbor-joining tree constructed from raw genetic distances representing relationships between samples. 170,000 SNPs were genotyped in 46 diverse dog breeds plus wolves using the Canine HD array. The Boxer branches are longer, which likely represent the influence of ascertainment bias, as the SNPs were discovered from sequence alignments involving the Boxer reference sequence. Originally published in Vaysse et al. (2011)

9.1 Head Phenotype

Brachycephaly is a phenotype resulting in a dramatic decrease in muzzle length accompanied by decreased length of the related bones (Fig. 10). Additionally, brachycephalic dog breeds, such as the Boxer, Bulldog, French Bulldog, and Pekingese, have slightly shortened and widened skulls.

An "across-breed" study was designed to investigate the genetic basis of the brachycephalic phenotype. This genome-wide association study design required control breeds lacking the brachycephalic phenotype and included dolichocephalic (long muzzle) and mesaticephalic (intermediate muzzle length) breeds. The dolichocephalic and mesaticephalic breeds included Akitas, Belgian Tervurens, Black

Fig. 9 Morphological variation in the dog. Dog breeds display extremes of morphological variation including body size and proportion, head size and shape, coat texture, color, and patterning. Clockwise from the left: the Bloodhound, the Chinese crested, the Dandie Dinmont Terrier, the Scottish Deerhound, the long-haired Chihuahua, and the French Bulldog (Original Image: Mary Bloom, American Kennel Club). Originally published in Shearin and Ostrander (2010)

Russian Terriers, Bloodhounds, Dalmatians, German Shepherds, and Great Danes. Bannasch et al. (2010) identified the location of the dog genomic region responsible for the brachycephalic phenotype using an across-breed genome-wide association approach. Using the Affymetrix Version 2 Custom Canine SNP arrays to generate genotype calls, the authors successfully identified a brachycephalic head locus that mapped to a region of chromosome 1 between 59.5 and 59.8 Mb (Bannasch et al. 2010). To more clearly delineate the region of association, the investigators used 88 affected dogs and 185 unaffected dogs to genotype a set of 49 SNPs overlapping the most significantly associated region of the originally identified interval. The results of this genotyping revealed a smaller 31 kb genomic interval that overlapped a homozygous haplotype encoding a single gene, THBS2 within brachycephalic breeds (Bannasch et al. 2010).

Schoenebeck et al. (2012) searched for additional genes modulating the multigenic phenotype and cranioskeletal features differentiating dolichocephalic skulls from brachycephalic skulls. In order to more completely characterize the anatomical and geometric differences associated with phenotypic variation in canine skull shape, the authors digitally captured 51 stereotyped anatomical landmarks from 533 skulls obtained from museums representing 120 breeds and 4 gray

Fig. 10 Brachycephaly in dogs. Comparison of photographs (Photos Mary Bloom, courtesy of AKC) and skulls from a German Shepherd dog with a wild-type skull shape (non-brachycephalic) and a brachycephalic Boxer. Originally published in Bannasch et al. (2010)
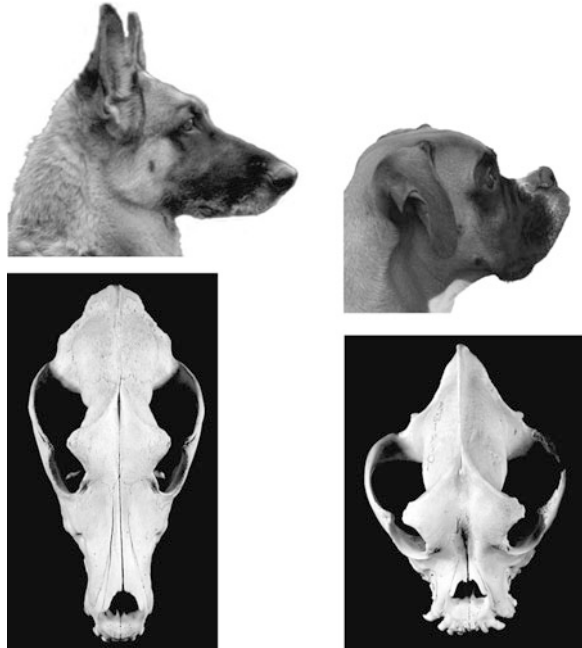
wolf subspecies. The variance captured in Principal Component 1 (PC1) (59% variation) corresponds to anatomical differences in rostrum length and angle, palate and zygomatic arch width, and depth of neurocranium, which comprise the cranioskeletal features giving rise to either a brachycephalic or dolichocephalic skull phenotype (Fig. 11).

Schoenebeck et al. (2012) used one set of breed samples for phenotypic measurements (the museum specimens) and another set of breed samples (the DNA samples) for genotyping because purebred dogs conform to a specific morphological standard that is shared among members of the breed. Morphological phenotypes, such as skull shape, are uniformly constrained by the breed. Strong genotype associations were found with PC1 (i.e., variations in skull morphology differentiating brachycephalic skull phenotype from dolichocephalic skull phenotype) associated with polymorphic markers located at specific locations within domestic dog, *Canis familiaris*, chromosomes (denoted CFA): CFA5.36476657, CFA24.26359293, CFA30.35656568, and CFA32.8384767. Some additional markers were weakly implicated on CFA9, CFA13, and CFA30 and another one on CFAX (Schoenebeck et al. 2012).

Schoenebeck et al. (2012) reasoned that skull shape variation is a consequence of artificial selection, and therefore they expected the major loci to exhibit reduced observed heterozygosity (H_o) and elevated genetic differentiation (F_{ST}), both of which are strong indicators of selective sweeps. The CFA32 quantitative trait locus (QTL) was selected as a major focus because it was in the top 2 most associated non-allometric loci that showed strong evidence of selection. The shared
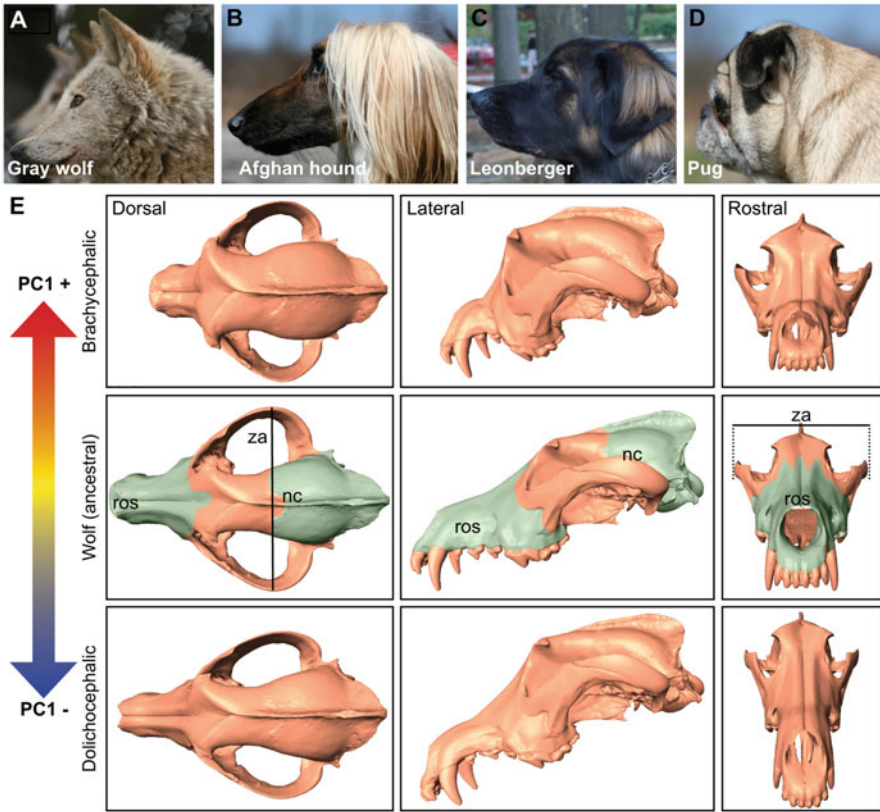
Fig. 11 Quantitative and qualitative assessments of PC1 on canine cranioskeletal shape. (**a**) Gray wolf (mesocephalic, ancestor to dogs) (**b**) Afghan hound (dolichocephalic), (**c**) Leonberger (mesocephalic), (**d**) Pug (brachycephalic). (**e**) Surface scans of a gray wolf skull illustrate morphological changes associated with PC1. Columns (left to right) are dorsal, lateral, and rostral views. Top row: a gray wolf skull morphed by positive PC1. Middle row: a gray wolf skull (no morphing). Bottom row: a gray wolf skull morphed by negative PC1. Pseudocoloring of the gray wolf skull indicates rostrum (ros) and neurocranium (nc). Line indicates width of the zygomatic arches (za). Originally published in Schoenebeck et al. (2012)

haplotypes for CFA32 QTL among six of the seven most brachycephalic breeds (Boston Terrier, Bulldog, Brussels Griffon, French Bulldog, Pekingese, and Pug) defined a 190 kb genomic region in between 8.15 and 8.34 Mb, within which two genes (*PRKG2* and *BMP3*) were located (Schoenebeck et al. 2012).

In order to ascertain genotype-phenotype association within this interval, Schoenebeck et al. (2012) performed whole genome sequence survey from 11 dogs of diverse skull phenotype (including the brachycephalic breeds of Bulldog and Pekingese breeds). The authors identified the SNP at position 8,196,098 that causes a missense mutation in *BMP3* in which a phenylalanine is changed into a leucine (F452L mutation). The substitution of leucine in place of phenylalanine was

predicted to be disruptive to the *BMP3* functional structure. Upon comprehensive genotyping of 842 dogs across 113 breeds, it was found that the F452L mutation is almost always fixed in brachycephalic breeds. This suggests that the missense polymorphism may be the underlying cause for the brachycephalic phenotype (Schoenebeck et al. 2012).

9.2 Genomic Basis of Breed-Associated Morphological Traits

A fundamental question in canine genomics is "*What genomic mechanism enabled selective breeding to produce the tremendous diversity of morphological phenotypes observed in present day dog breeds*?" Boyko et al. (2010) addressed this question by genome-wide scans of SNP variation and genome-wide association mapping of morphological traits using 60,968 SNP genotypes of 915 dogs covering 80 domestic dog breeds coupled with 83 wild canids and 10 outbred African shelter dogs. The genotype map was combined with external measurements using breed standards, museum specimens, and individual dogs to identify genomic regions associated with breed-specific phenotypic variation among 57 morphological traits. One of the purposes of the study was to assess whether most breed-associated phenotypic variation is the consequence of large-effect QTLs or whether most of the observed phenotypic differences arise via the action of many QTLs of relatively weak effects. The answer to this question will provide a better understanding of how domestication and artificial selection have impacted the dog's genome (Boyko et al. 2010).

Boyko et al. (2010) performed a genome-wide scan to detect signatures of recent selection and allele sharing between dog breeds. Because the data supports the idea that relatively little sharing of IBD segments occurs among individuals from different breeds, it is reasonable to expect that when coincident sharing occurs between breeds with a similar phenotype, the shared genomic segments are likely encoding the genetic variation for that trait. The top 11 most extreme F_{ST} regions of the dog genome contained SNPs with $F_{ST} \geq 0.57$ and having a minor allele frequency (MAF) ≥ 0.15 (Boyko et al. 2010). Among the 11 regions detected with high F_{ST}, 6 are tightly linked to genetic variation known to affect canine morphological phenotypes. For example, the 167 bp insertion in *RSPO2* was associated with the fur growth and texture phenotype; the IGF1 haplotype was associated with small body size; an inserted retrogene (*Fgf4*) was associated with short limb length; and three genes modulating coat color phenotypes in dogs were also associated with the identified intervals *ASIP*, *MC1R*, and *MITF*. Additional regions with high F_{ST} were identified: CFA10.11465975 (associated with body weight) and CFA1.97045173 (associated with muzzle length) (Boyko et al. 2010).

Boyko et al. (2010) performed the genome-wide association scans by measuring 55 morphological parameters in order to identify genotype-phenotype associations, especially morphological traits that vary between dog breeds. Additionally, the authors included genomic regions contributing to variation in body size (variation is greatest across dog breeds than any other terrestrial species) as well as ear
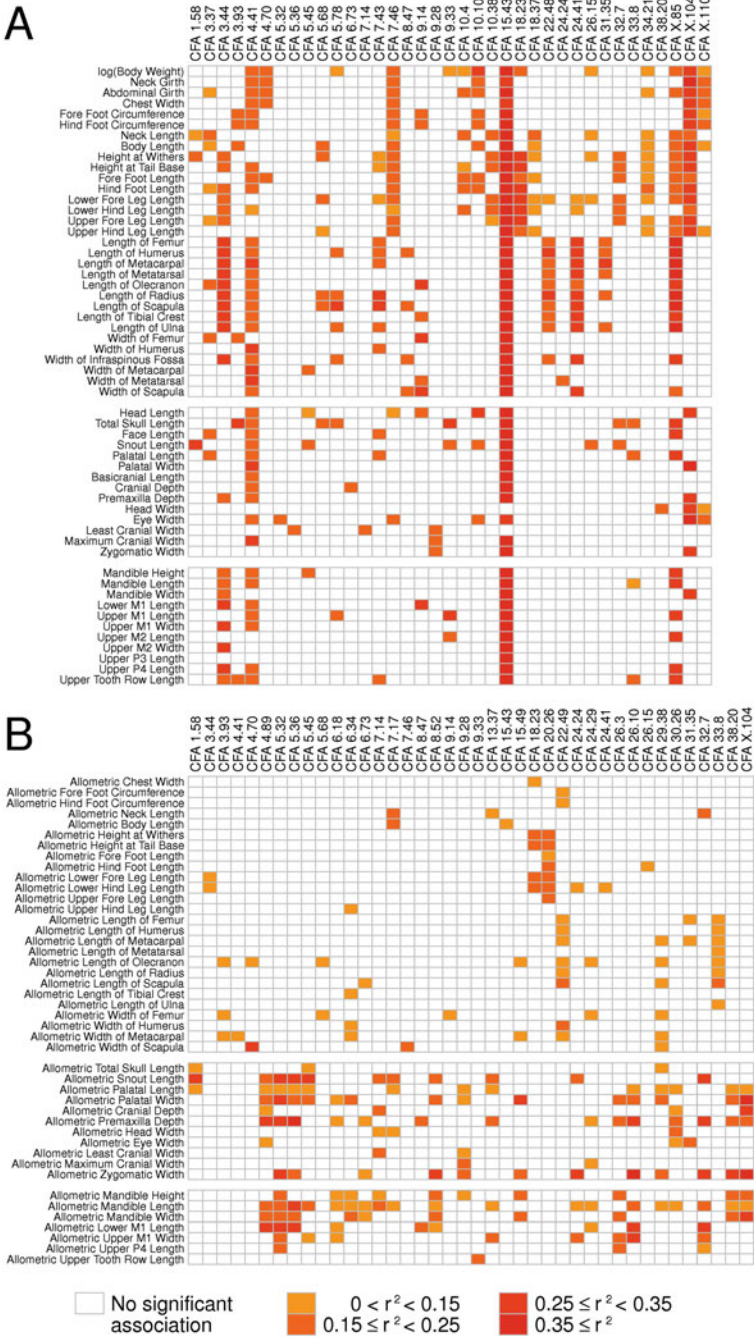
Fig. 12 Summary of associations across genomic regions for multiple traits. Each row corresponds to a trait [either (**a**) absolute or (**b**) proportional], and each column corresponds to a genomic region that has been found associated with at least one trait. The shading of each rectangle shows the R^2

floppiness. The genomic scan for body size [where body size = log (body weight)] resulted in the identification of multiple significant associations. The six strongest signals occurred at CFA15.44226659, CFAX.106866624, CFA10.11440860, CFAX.86813164, CFA4.42351982, and CFA7.46842856. Interestingly, the first four signals identified in the body size variation scan correspond to some of the highest F_{ST} values identified in the genome, along with CFA4 which has an $F_{ST} = 0.46$, consistent with diversifying selection among breeds for body size. Interestingly, in all six regions, wolves are not highly polymorphic (MAF < 0.1), and except for the CFA10 signal, the derived allele is at highest frequency in small breeds (Boyko et al. 2010).

Another trait that exhibits considerable variation across breeds is ear type. All adult wild canids have erect ears, yet dog breeds are fixed for a variety of ear positions including floppy ears. This juvenile type trait is retained by adults of certain breeds in a variety of domesticated mammals, such as dogs, cattle, goats, and rabbits. SNPs associated with breeds fixed for erect or floppy ears were identified and shown to be associated with a single interval on CFA10 that may underlie the ear position phenotype (Boyko et al. 2010). A third trait of interest in the Boyko et al. (2010) study was muzzle length, which varies tremendously across dog breeds. Similar to floppy ears, short snout is another paedomorphic trait. The strongest association signals were CFA1.59832965 and CF5.32359028, having F_{ST} values of 0.55 and 0.42, respectively. These polymorphisms are only found in brachycephalic breeds at high allele frequency (Boyko et al. 2010).

Boyko et al. (2010) constructed a multi-SNP predictive model for each trait. For the models of body weight, ear type, as well as most of the measured traits, the majority of the breed-associated variance was observed in fewer than four loci (Fig. 12). Correlated traits, such as femur length and humerus length, exhibited similar SNP associations. For the set of 55 measured traits, the average proportion of variance explained by the top 1, 2, and 3 SNP models was $R^2 = 0.52$, 0.63, and 0.67, respectively. The authors made the case that, after controlling for body size, mean proportion of variance explained by the models was still considerable, with $R^2 = 0.21$, 0.32, and 0.4, respectively. It is worth mentioning that the most significant genomic regions were similar even using naïve association scans that did not control for population structure. In terms of breed mapping, relatively little population structure was shared among the breeds. Subsequently, whatever portion of the population structure, which might have been shared among the breeds, was small enough to avoid biasing the association inferences (Boyko et al. 2010).

Boyko et al. (2010) state, for the majority of traits investigated, that a few QTLs of large effect determined the phenotype differences between breeds. These QTLs mapped to specific locations on *Canis familiaris* chromosomes (CFA). As an

---

Fig. 12 (continued) statistic of the single marker model for the trait for all significant associations ($p < 5.0e{-}5$ for absolute external traits, $p < 1.0e{-}4$ for skeletal and proportional traits after correcting for population structure). When multiple SNPs in the region are significant, the largest value of the R^2 statistics is reported. Originally published in Boyko et al. (2010)

example, they site the proportional height at withers for which they identified a large-effect QTL on chromosome CFA18, where they had previously determined a *fgf4* retrogene that confers the phenotype associated with the chondrodysplasia dispro-portional dwarfism in Basset Hounds, Corgis, and Dachshunds. Similarly, skull shapes were largely determined by genomic regions on CFA1, CFA5, CFA26, and CFA32, along with CFAX.105274087–106866624 region (also associated with body size). Most of these regions were also associated with dental phenotypes along with a strong association on CAF16. It seems that the relationship between phenotypes and associated genomic intervals in domestic dog breeds can be best described as a set of related phenotypes under the direct control of a few genomic regions (Boyko et al. 2010).

10 Genes, Mutations, and Genomic Regions Contributing to Clinically Relevant Phenotypes (Disease Conditions) in Dog Breeds

In the past few years, new advances have been achieved using genome-wide association studies (GWAS) and high throughput sequencing to unveil novel muta-tions in dog populations associated with clinically relevant phenotypes. These phenotypes span numerous organs, cell types, and body systems. Some interesting examples across a variety of body systems and dog breeds are described below.

10.1 Cardiovascular

Cardiovascular disease affects different dog breeds including the Newfoundland, Whippet, and Doberman Pinscher. Mitral valve degeneration is the most prevalent type of heart disease in dogs and is acquired during aging as degenerative lesions accumulate on the mitral valve. Over time, these lesions result in abnormal valve morphology and function. In severe cases, the mitral valve may prolapse and cause undesirable phenotypes, such as mitral regurgitation and left-sided congestive heart failure.

Stern et al. (2015) used the 170,000 canine high-density (HD) genotyping SNP chip and identified a region in the vicinity of position 57,770,326 on canine chromosome 15, which is near the interval of 58,506,916 and 60,140,841 that was also associated with mitral valve disease compared to normal dogs lacking evidence of mitral valve disease. Within this region is follistatin-related protein 5 precursor as well as some other genes including neuropeptide Y receptors. A region on chromo-some 2 also exhibited partial evidence of association peaking at 37,628,875 which is in proximity to rho GTPase-activating protein 26. In the discussion, the authors implicate follistatin-related protein 5 with another gene (*WFIKKN2*) that is involved

in metalloproteinase inhibition activity. Because metalloproteinase activity has been considered a part of the mitral valve disease pathophysiological mechanisms, these two genes represent viable candidates for the undesirable clinical trait of mitral valve disease (Stern et al. 2015).

The Doberman Pinscher is one of the most commonly reported canine breeds with familial dilated cardiomyopathy, which has been linked to congestive heart failure and sudden cardiac death. Meurs et al. (2012) performed a GWAS using a commercial "Canine Genome Array" containing 49,663 SNP markers and identified a locus on CFA14 (Meurs et al. 2012). Fine-mapping of additional SNPs localized a potential haplotype at 23,774,190–23,781,919 region from the same chromosome. DNA sequencing identified a 16 bp deletion in the 5′ donor splice site of intron 10 from the gene encoding the mitochondrial pyruvate dehydrogenase kinase 4 (*PDK4*) in affected dogs. The authors next demonstrated that *PDK4* transcripts derived from the homozygous deletion genotype exhibit decreased expression of exons 10 and 11. This study tested 232 animals, with 66 affected and 66 unaffected Doberman Pinschers, plus 100 healthy dogs from 11 other breeds. The target mutation was identified in 54 out of 66 affected dogs (82%, with 45 heterozygotes and 9 homozygotes) and 26 out of 66 of unaffected dogs (39%, with 18 heterozygotes and 8 homozygotes). Some of the 100 unaffected dogs, representing 11 other breeds, appeared to show the mutated allele as well. Electron microscopy of myocardium from affected dogs demonstrated several mitochondrial disorganization features, suggesting a dysfunction of *PDK4* enzyme due to the mutation (Meurs et al. 2012). The fact that the presence of an associated allele may not always correlate with the associated phenotype underscores the complexity of genetics.

The Irish Wolfhound is another breed that is predisposed to cardiac disease, specifically dilated cardiomyopathy, with up to 20% of dogs in the breed exhibiting the undesirable clinical phenotype. Philipp et al. (2012) performed a genome-wide association study using 190 Irish Wolfhounds. Dilated cardiomyopathy phenotypes were diagnosed with echocardiographic exams. Control dogs were at least 7 years old with no signs of the dilated phenotype. The authors identified six loci corresponding to CFA1 at 123,630,555; CFA10 at 24,159,608 (*ARHGAP8* gene); CFA15 at 61,260,406 (*FSTL5* gene); CFA17 at 58,604,566; CFA21 at 40,670,543 (*PDE3B* gene); and CFA37 at position 31,801,266. The authors report that their associated regions overlapped with genes known to cause dilated cardiomyopathy in humans (Philipp et al. 2012). The human form of dilated cardiomyopathy is a cause for heart transplants, and in the absence of transplantation, chronic heart failure can occur. About half of human cases are inherited, and more than 60 genes have been implicated in the pathology (Toro et al. 2016).

In another example of cardiovascular phenotypes in dogs, Stern et al. (2014) used a pedigree analysis of 45 Newfoundlands, of which 9 exhibited a subvalvular aortic stenosis (SAS) phenotype. Twelve additional dogs in the pedigree displayed systolic heart murmur phenotypes along with either evidence of aortic insufficiency or a subvalvular ridge or both. When dogs with the aortic insufficiency and/or subvalvular ridge phenotypes were bred to normal dogs, offspring displayed undesirable

cardiac phenotypes. A genome-wide association study followed by genomic sequencing identified a mutation in the exonic region of the phosphatidylinositol-binding clathrin assembly protein gene (*PICALM*). Interestingly, *PICALM* is involved in morphogenesis of the heart. Stern et al. (2014) report that the phenotype is likely caused by a 3 bp exonic insertion in the *PICALM* (599K_600LinsL mutation) that was detected and associated with the development of SAS in that breed. Immunohistochemistry validated the presence of *PICALM* protein in the canine myocardium and area of the subvalvular ridge. Overall, 96.1% of the SAS-affected Newfoundland dogs displayed the codon insertion mutation (34.6% homozygous and 61.5% heterozygous), while only 26% of non-affected ones possessed the mutation (4.3% homozygous and 21.7% heterozygous). The authors state that none of 180 control dogs of 30 different breeds possessed the mutation in any form (Stern et al. 2014).

Following the report by Stern et al. in 2014, Drogemuller et al. (2015) provided evidence suggesting that the mutation reported by Stern et al. may not in fact be the causative allele associated with subvalvular aortic stenosis in Newfoundlands. Among the evidence presented, Drogemuller et al. (2015) question the experimental design that was used, pointing out that (a) the number of cases and controls used in the association study would not provide the expected power needed to identify a locus associated with a nondominant mode of inheritance (Drogemuller et al. 2015). Furthermore, Drogemuller et al. (2015) report a replication of portions of the original study and fail to reproduce the findings reported by Stern et al. (2014).

10.2 Endocrinology

An endocrine phenotype of clinical interest is obesity. Obesity and greater food motivation were found as a genetic predisposed disorder in Labrador retrievers (Raffan et al. 2016). The associated gene is pro-opiomelanocortin (*POMC*) that encodes a pro-protein which is cleaved into several bioactive peptides, including b-MSH (melanocyte-stimulating hormone) and b-endorphin. The associated genotype is a 14 bp deletion responsible for a frameshift after the glutamate at the position 188 (p.E188fs). It is predicted to disrupt the coding sequence of *POMC* and cause loss of production of b-MSH and b-endorphin which results in increased body weight with a mean effect size of 1.90 kg per deletion allele. Therefore, it indicates a dominant dosage effect trait. Adiposity and food motivation were polymorphism associated phenotypes in both Labrador Retrievers and the closely related Flat-Coat Retrievers (FCRs). The mutation is significantly more common in Labradors selected to become assistance dog breeding stock (allele frequency = 0.45) than those selected to be companions (allele frequency = 0.12) (Raffan et al. 2016). In humans, *POMC* mutations that produced aberrant forms of b-MSH reveal that this is an important hormone in controlling appetite and obesity development (Challis et al. 2002; Lee et al. 2006). Mice selectively lacking b-endorphin are hyperphagic and obese (Appleyard et al. 2003). Taken together, these findings suggest that the

loss of both neuropeptides in dogs carrying *POMC* p.E188fs could contribute to the observed obese phenotype.

10.3 Ophthalmology

Progressive retinal atrophy (PRA) is a group of inherited eye diseases characterized by retinal degeneration that culminates to blindness in dogs and is often described as the equivalent of retinitis pigmentosa (RP) in humans. It is noteworthy that PRA in dogs has been reported in over 100 dog breeds. Three studies, two in Golden Retrievers and one in Shelties, have uncovered three PRA-related genes. The first study in Golden Retrievers leveraged a genome-wide association study design to ultimately identify a frameshift mutation within the canine solute carrier gene *SLC4A3*. The undesirable allele was present in 56% of PRA dogs and exhibited recessive inheritance with 100% penetrance (Downs et al. 2011).

The second study in Golden Retrievers (GRs), used GWAS in 10 PRA cases and 16 controls, identifying an association of a 737 kb chromosome 8 (CFA8) locus containing six genes with a clinical ocular phenotype. Two of the genes (*TTC8* and *SPATA7*) have already been described as RP-associated in humans. *TTC8* encodes a protein that is a part of the BBSome complex which is responsible for ciliary membrane biogenesis. Affected dogs showed a single nucleotide deletion in *TTC8* exon 8. The frameshift mutation is predicted to cause a premature stop codon. In the investigated cohort, this genotype (*TTC8* c.669delA) is recessive, segregating correctly in 75.9% of the tested cases (22/29), whereas none of the PRA controls are homozygous for the mutation, only 3.5% carry the PRA-associated allele, and 96.5% are homozygous wild type (Downs et al. 2014).

Identifying genes associated with PRA provides a mechanism for developing breeding programs that can eventually remove these undesirable alleles from affected breeds. The pathophysiology and clinical progression of PRA have been well characterized within the Swedish Vallhund dogs by Cooper et al. (2014). A third study reported by Wiik et al. (2015) identified the *CNGA1* gene on CFA13 as a novel PRA-related locus using a genome-wide association approach with 15 Shetland Sheepdog (Sheltie) cases and 14 controls. *CNGA1* is also known to be involved in human RP. This gene encodes a protein involved in phototransduction, by forming cGMP-gated cation channel in the plasma membrane that allows depolarization of rod photoreceptors. Sequencing of this gene in affected Shelties identified a 4 bp deletion in exon 9 (c.1752_1755delAACT). Similar to the *TTC8* mutation in Golden Retrievers, *CGNA1* also alters the translation frame and generates a truncated protein caused by premature termination codon (Wiik et al. 2015).

Besides PRA, other ocular phenotypes affect dogs, such as glaucoma. Two metalloprotease genes *ADAMTS10* and *ADAMTS17* are implicated in primary open angle glaucoma (POAG) in dogs: the former in Beagle (Kuchtey et al. 2013) and Norwegian Elkhound breeds (Ahonen et al. 2014) and the latter in Basset Hound and Basset Fauve de Bretagne breeds (Oliver et al. 2015). Regarding the latter study, 226 Basset Hounds and 27 Basset Fauve de Bretagne dogs were provided an

ophthalmic examination and diagnosed for POAG. The affected Basset Hounds displayed homozygosity for a 19 bp deletion in *ADAMTS17* exon 2 that leads to a frameshift predicted to form a truncated protein. Fifty clinically unaffected Basset Hounds were genotyped for this mutation as either heterozygous or homozygous for the wild-type allele. The affected Basset Fauve de Bretagne dogs contained a nonsynonymous substitution in *ADAMTS17* exon 11 causing a glycine to serine amino acid exchange (G519S) in the disintegrin-like domain that might be related to protein dysfunction. Unaffected Basset Fauve de Bretagne dogs were either hetero- zygous for the mutation (5/24) or homozygous for the wild-type allele (19/24). Therefore, evidence suggests that both independent POAG-associated mutations are recessive in the two different breeds examined (Oliver et al. 2015).

10.4 Craniofacial

Wolf et al. (2015) described a mutation on the dog's chromosome 27, encoding a frameshift mutation within the *ADAMTS20* metallopeptidase gene (c.1360_1361delAA or p.Lys453Ilefs*3), that leads to a cleft lip with or without cleft palate (CL/P) phenotype in the Nova Scotia Duck Tolling Retriever (NSDTR). This undesirable phenotype exhibits a recessive mode of inheritance (Wolf et al. 2015). CL/P is the most commonly occurring craniofacial congenital disorder. Interestingly, the same study that found *ADAMTS20* as the CL/P-target gene in NSDTR dogs has also reported a suggestive association of the same gene to CL/P human cases in a family-based association analysis (DFAM) using a Guatemalan cohort composed of 25 CL/P phenotypes, 420 unaffected relatives, and 392 controls. In dogs, the mutation alters the reading frame and generates a premature stop codon within the metalloprotease domain of *ADAMTS20* protein. In humans it seems to be associated with the SNP rs10785430 within *ADAMTS20*, but further studies are required to assure whether it alters the protein function.

10.5 Dermatology

Canine atopic dermatitis (CAD) is a chronic inflammatory skin disease triggered by environmental allergens that react with epithelial and immune cells. GWAS and fine-mapping analyses revealed a 9-SNP-containing haplotype overlapping *PKP2* gene that predisposes German Shepherd dogs to CAD. PKP2 encodes plakophilin-2 protein, which is involved in the synthesis of desmosomes, a cell adhesion structure (Tengvall et al. 2016). The haplotype spans ~280 kb on chromosome 27 (CFA27) which encompasses a rare ~48 kb locus shared only with other high-risk CAD breeds. Transient transfections followed by luciferase reporter assays indicated that seven out of the nine CAD-associated SNPs within that haplotype appeared to have enhancer activity with allelic differences in either epithelial or immune cells. These cells include Madin-Darby canine epithelial cell line from Cocker Spaniel (MDCK),

human keratinocyte cell line (HaCaT), human T cell line (Jurkat), and human erythromyeloblastoid leukemia cell line (K562). A top SNP (CFA27:19,086,778) displayed high activity in keratinocytes with 11-fold induction of luciferase transcription by the risk allele (T/T) versus 8-fold by the control allele (C/C) ($p = 0.003$). It also mapped close (∼3 kb) to an ENCODE skin-specific enhancer region. Those experiments suggest that GSDs' predisposition to CAD is associated with multiple variants combined in a risk haplotype that may contribute to an altered expression of the PKP2 gene in immune and epithelial cells (Tengvall et al. 2016).

10.6 Pigmentation

A recessive genotype, within the solute carrier family 45, member 2 gene (*SLC45A2*), is responsible for albinism in dogs. The *SLC45A2* protein is found in melanocytes, and, although its exact function is still being studied, it is likely to be involved in melanin synthesis. A large deletion (g.27,141_31,223del) in *SLC45A2* was associated with oculocutaneous albinism (OCA) in Doberman Pinschers (Fig. 13) that were homozygous for that mutation, whereas the albino Lhasa Apso showed homozygosity for a nonsynonymous substitution in the seventh exon of *SLC45A2* (c.1478G > A) that resulted in a switch from glycine to aspartate (p.G493D) (Wijesena and Schmutz 2015). This same study revealed that an albino Pekingese, two albino Pomeranians, and one albino mixed breed dog that was small and long-haired were also homozygous for the 493D allele. Colored offspring from those small long-haired albinos were heterozygous for this allele, clearly indicating that it is a recessive genetic trait. Structural bioinformatics investigation has predicted that the 11th transmembrane domain (where the 493rd amino acid is located) from the *SLC45A2* (p.G493D) protein has an altered structure, which might be deleterious for the proper protein function and, consequently, leads to the albino phenotype due to the lack of melanin production. However, an albino Pug was genotyped as homozygous for the 493G allele, indicating that although 493D allele is related to albinism in some small, long-haired dog breeds, it does not explain all albinism in dogs (Wijesena and Schmutz 2015).

10.7 Musculoskeletal

Mosher et al. (2007) identified the myostatin gene as the cause of increased muscle mass in Whippets. Interestingly, Whippets, like Greyhounds, are bred for racing. The Whippet is a small dog breed weighing approximately 9 kg. Within the population of race-bred Whippets, a "Bully Whippet" phenotype emerged in which heavily muscled Whippets were produced by breeders (Fig. 14). Although owners report that the Bully Whippets are healthy with some incidents of muscle cramping, they are never the less euthanized as they do not conform to the breed standard. The authors report that a total of 22 Whippets were sequenced across the
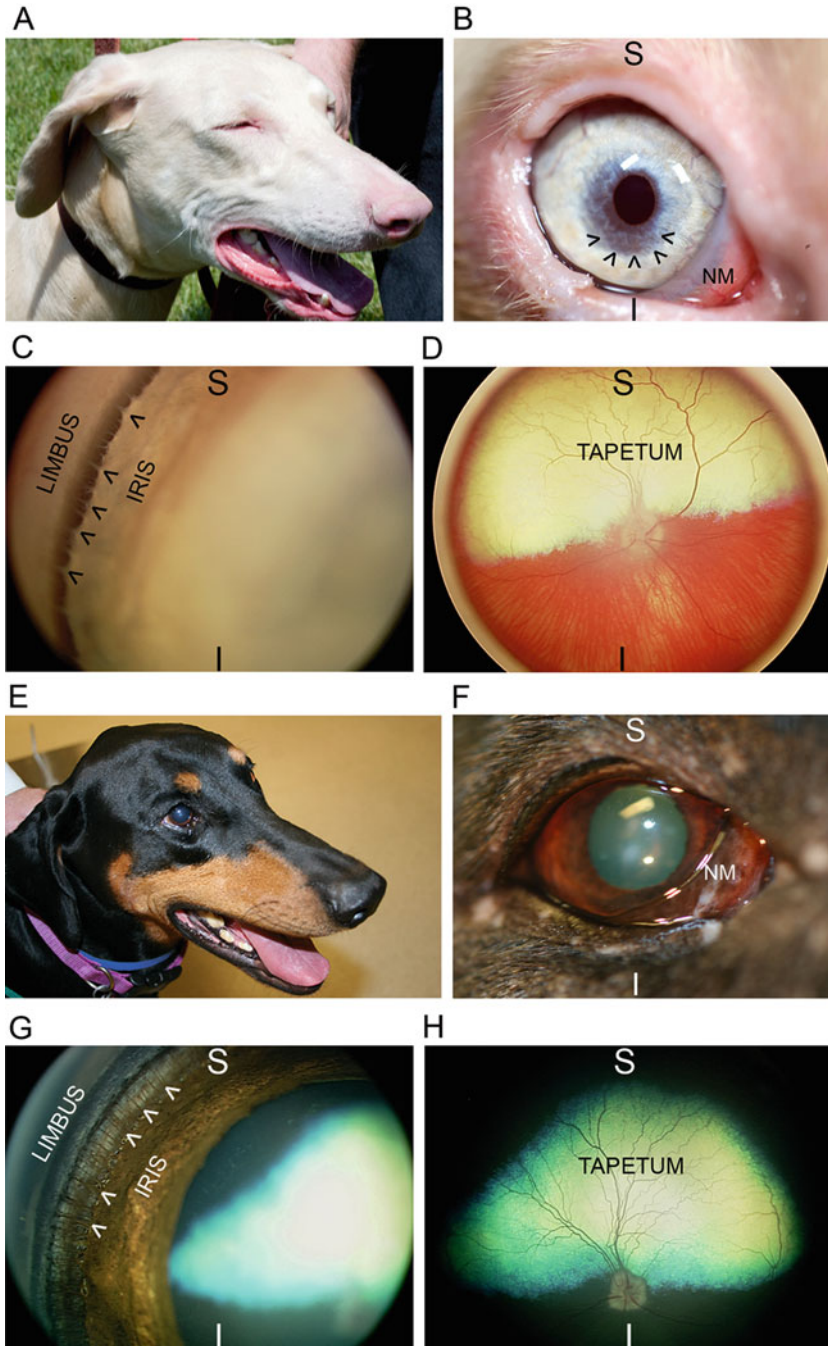
Fig. 13 Ocular phenotype of white Doberman Pinschers. Images taken from white Doberman Pinschers (top row) and black standard-color Doberman Pinscher (bottom row). An image of white Doberman Pinscher head (**a**) demonstrates lightly pigmented nose, lips, and eyelid margins

three exons and most of the introns in the myostatin gene. Among those sequenced, all four with the Bully Whippet phenotype were homozygous for a 2-bp deletion within the third exon that removes nucleotides 939 and 940 resulting in a premature stop codon. Of the five dogs that sired or whelped a Bully Whippet, all were heterozygous for the 2-bp deletion mutation. None of the remaining 13 Whippets, which all lacked the bully phenotype and had no familial history of the phenotype, carried the 2-bp deletion mutation (Mosher et al. 2007).

The authors determined that the bully phenotype displayed a simple autosomal mode of inheritance. Furthermore Mosher et al. (2007) provided statistical support for the idea that heterozygous Whippets contain, on average, 17% more mass per centimeter of height compared to homozygous wild-type Whippets (p-value = 0.00017). When the authors analyzed the genotypes of 85 racing dogs, for which racing results were available, an association between the mutation and racing performance was detected. Specifically, among dogs that were heterozygous for the mutation ($N = 12$), 66% were classified as top racers, while less than 17% of wild-type dogs received the same top ranking ($n = 72$). The Bully Whippets are too heavily muscled to perform well in races, while the heterozygotes exhibit ideal racing performance associated with lean muscle. The authors ultimately sequenced 15 different breeds and determined the haplotypes spanning the myostatin gene (Mosher et al. 2007).

10.8 Neoplasia

Cancers are genetically inherited diseases that occur in multiple species including dogs and humans. Identifying tumorigenesis-associated mutations is of great importance in veterinary medicine; dog's neoplasias are also valuable spontaneous

Fig. 13 (continued) compared with the same darkly pigmented structures in SDP (**e**). A closeup image of WDP eye (**b**) shows nonpigmented leading edge of the nictitating membrane (NM), tan-colored iris base transitioning to blue at pupillary margin, and oval-shaped dyscoric pupil aperture. The black arrowheads (in **b**) demarcate a region of significant iridal stromal thinning that was noted on examination to transilluminate (not shown in image) with retroillumination by light reflected from the tapetum lucidum. SDP eye (**f**) shows darkly pigmented margin of the nictitating membrane (NM) and brown iris with a round pupil aperture. WDP gonioscopy image (**c**), which allows visualization of structures lying within the iridocorneal angle (in images **c** and **g**, this region lies between the words "LIMBUS" and "IRIS"), shows that fibers of the pectinate ligament (demarcated by black arrowheads) are of a similar tan color to the iris base, whereas fibers of the pectinate ligament (demarcated by white arrowheads) are dark brown in SDP (**g**). WDP fundus image (**d**) shows yellow-colored tapetum lucidum (labeled "TAPETUM") and significant hypopigmentation of the retinal pigment epithelium and choroid allowing visualization of the choroidal vasculature. SDP fundus image (**h**) shows green-colored tapetum lucidum (labeled "TAPETUM") and heavy pigmentation of the non-tapetal fundus. For orientation purposes, images taken at higher magnification (**b–d** and **f–h**) have the superior (S) and inferior (I) globe positions labeled. Originally published in Winkler et al. (2014)
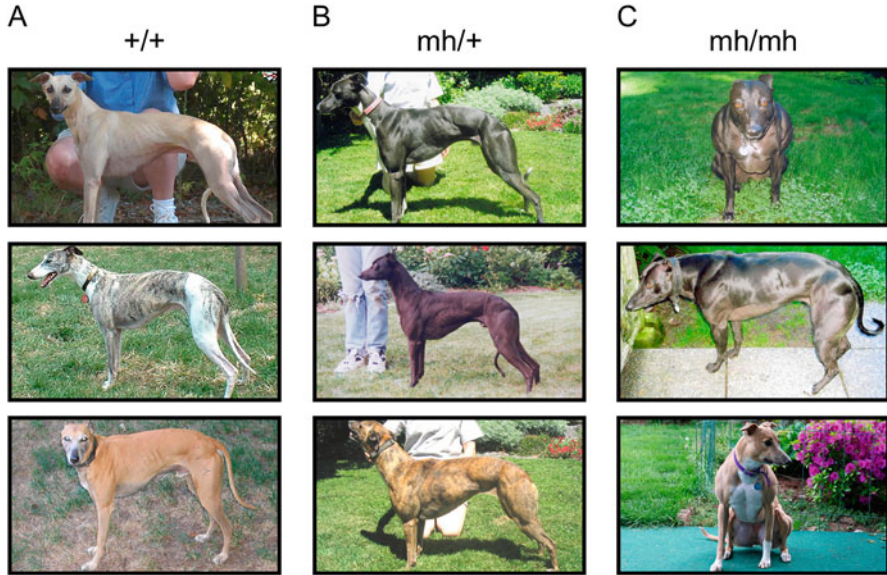
Fig. 14 Whippets with each of the three potential myostatin genotypes. (**a**) Dogs have two copies of the wild-type allele (+/+). (**b**) Dogs are heterozygous with one wild-type allele and one mutant cys → stop allele (mh/+). (**c**) Dogs are homozygous for the mutant allele with two copies of the cys → stop mutation (mh/mh). All photos represent unique individuals except for the top and middle panels in the right-hand column. Originally published in Mosher et al. (2007)

models for better understanding human cancer. The same GWAS approach can also be applied in cancer. For instance, a GWAS containing 39 dog glioma cases and 141 controls from 25 dog breeds identified a significant locus on chromosome 26 (CFA26) (Truvé et al. 2016). Resequencing of a 3.4 Mb target region was performed, revealing 56 SNPs that best fit the association pattern between the resequenced cases and controls. Three candidate genes were highly associated with glioma susceptibility: a calcium-/calmodulin-dependent protein kinase 2 (*CAMKK2*), a P2X ligand-gated ion channel 7 (*P2RX7*), and an mRNA translation reinitiation factor (*DENR*) that influences the migration of cerebral cortical neurons in mice (Haas et al. 2016).

Similarly, an investigation into canine mast cell tumors (CMCT) made use of GWAS in Golden Retrievers from two continents [127 from the United States (70 cases and 57 controls) and 146 from Europe (71 cases and 75 controls)], identifying different regions in the genome associated with risk of CMCT in the two populations (Arendt et al. 2015). Sequencing of GWAS-rescued regions and subsequent fine-mapping identified a *GNAI2* SNP associated with development of CMCT. The *GNAI2* gene encodes an alpha subunit of guanine nucleotide-binding proteins (G proteins) that are transducers in various transmembrane signaling systems and play a role in cell division. The identified SNP introduces an alternative splice form that gives rise to a truncated protein. In addition, CMCT-associated

haplotypes harboring the hyaluronidase genes *HYAL4, SPAM1*, and *HYALP1* on CFA14 and *HYAL1, HYAL2*, and *HYAL3* on CFA20 were identified as separate risk factors in US GRs and European GRs. This suggests that turnover of hyaluronic acid is important for the development of CMCT (Arendt et al. 2015).

It appears that tumorigenesis and cancer associated phenotypes may arise through a variety of mechanisms within the dog. Borge et al. (2015) assessed copy number variations using microarrays to assess genotypes within 117 canine mammary tumors obtained from 69 dogs. The authors point out that cancer cell genomes differ from the host genome through single nucleotide polymorphisms, gain/loss of large chromosomal regions via duplication/deletion of large genomic segments, and expanded/contracted copy numbers of certain loci. Borge et al. (2015) employed the Illumina 170 K canine HD array. Their analysis identified a number of genes with known cancer associations in humans that were frequently amplified or deleted in canine mammary tumors. Some of the genes frequently amplified in the tumors included *BCL6*, *FGFR2*, *MITF*, *MYC*, and *NPM1*, while genes exhibiting deletion loss within canine mammary tumors included *PTEN*, *BMPR1A*, *KDM5C*, *KDM6A*, and *PRF1* (Borge et al. 2015).

Squamous cell carcinoma of the digit (SCCD) in Standard Poodle (STPO) is a locally aggressive cancer that affects only dark coat color individuals. GWAS in 31 STPO SCCD cases and 34 unrelated black STPO controls detected a SNP peak on canine chromosome 15 (Karyadi et al. 2013). Fine-mapping pinpointed a region on the KIT Ligand (*KITLG*) locus. *KITLG* is a pleiotropic factor that acts in the development of both germ and neural cells as well as in hematopoiesis, which is involved in cell migration. Interestingly, the polymorphism within this locus implicated in modulating risk for squamous cell carcinoma appears to be a copy number variant within the transcriptional control region of the KIT locus that is predicted to contain regulatory enhancer elements (Karyadi et al. 2013).

Other mechanisms underlying susceptibility to cancer have been identified. Ferraresso et al. (2014) conducted an in-depth analysis of canine diffuse large B-cell lymphoma (DLBCL) and identified the downregulation of tissue factor pathway inhibitor 2 (*TFPI-2*) as a hallmark of lymph nodes associated with DLBCL. Moreover, the authors demonstrated that hypermethylation of the *TFPI-2* promoter, which increased as a function of age, correlated with decreased expression levels of the gene and demonstrated the age-dependent epigenetic alterations associated with canine DLBCL (Ferraresso et al. 2014).

Melin et al. (2016) performed a GWAS and identified three regions within the canine genome associated with mammary tumors in English Springer Spaniels. The study design consisted of 332 individuals, corresponding to 188 cases and 144 controls. The most significant genomic region was located on chromosome 11 and exhibited a complex architecture of numerous haplotypes spanning the centrosomal cell cycle regulator CDK5 regulatory subunit-associated protein 2 (*CDK5RAP2*). The genomic region spanned 700 kb and was refined to a smaller region of 446 kb. Within this region numerous SNPs, some of which are non-

synonymous and may alter protein function, were identified. Melin et al. (2016) assessed the relationship between the observed haplotypes using a phylogenetic tree approach and then calculated the frequency of cases and controls among the different haplotype groups. The cases within haplotype group 1 exhibited a lower frequency than in haplotype groups 2 and 3. The authors report that within this region of the genome, there are numerous noncoding RNAs such as miRNAs and snoRNAs, potentially implicating RNA-mediated interactions as contributing to mammary tumor susceptibility within this breed (Melin et al. 2016).

10.9 Many Clinically Relevant Traits in German Shepherd

Interestingly, the amount of genetic information about individual dog breeds is continuing to grow rapidly. The German Shepherd dog has been the focus of numerous genetic studies, and the results have opened the door to identification of genetic markers implicated in a significant number of phenotypes, many of which are associated with clinically relevant traits, such as atopic dermatitis and degenerative myelopathy (Table 1). Such knowledge provides opportunities for employing genotyping technology in the artificial selection of next-generation German Shepherds.

11 Conclusions and Future Perspectives

The tremendous wealth of dog genetics and genome information elucidated over the last couple of decades has dramatically altered our understanding of how the dog was domesticated and how artificial selection shaped it into the companion we live with today. There is no doubt that the 30,000 years of selective breeding have given rise to the dogs of today through the selection for specific traits that contribute to the dog's social fitness within human environments. Unfortunately, that same selection has contributed to undesirable clinical phenotypes in dogs as well. The tools of genomics have opened up the possibilities of inferring evolutionary history of dogs as well as the resulting impacts on the genome. Through the lens of genetics, we are able to discern exactly what biochemical molecules were altered in specific breeds during the domestication process. Furthermore, this window into the genome has allowed us to carefully begin to dissect the molecular events contributing to specific morphological phenotypes within particular breeds as well as the undesirable phenotypes associated with disease. These results, taken together, provide clear evidence that selection occurs in the presence of selective pressure and that artificial selection in dogs is an ongoing process. It is interesting to contemplate how dogs will continue to evolve in the future. No doubt it will be at the hands of humans; however, the tools available for aiding the artificial selection process are exponentially more powerful than they were during the original domestication and breed radiation

Table 1 Representative genetic markers associated with German Shepherd dog phenotypes

Phenotype	Gene	Genomic location	Polymorphism	Mode of inheritance	Reference
Activity-impulsivity	DRD4	Chromosome 18	Repeat polymor-phism in exon 3	Complex trait (genes and environment)	Hejjas et al. (2007)
Activity-impulsivity	TH	Chromosome 18	36 bp short allele versus 72 bp long allele	Complex trait (genes and environment)	Kubinyi et al. (2012)
Brown liver spots	TYRP1	Chromosome 11	Non-synonymous SNP	Autosomal recessive	Monteagudo and Tejedor (2015)
Canine anal furunculosis	ADAMTS16 and CTNND2	Chromosome 34 (both genes)	Multiple non-synonymous SNPs	Polygenic	Massey et al. (2014)
Canine atopic dermatitis (CAD)	PKP2	Chromosome 27	Multiple upstream regulatory SNPs	Unknown	Tengvall et al. (2016)
Canine atopic dermatitis (CAD)	PKP2	Chromosome 27	~209 kb region with 2 haplotype blocks	Unknown	Tengvall et al. (2013)
Canine degenerative myelopathy (CDM)	SOD1	Chromosome 31	Alters amino acid encoded at position 118	Autosomal recessive	Holder et al. (2016)
Canine inflammatory disease	NOD2	Chromosome 2	Four non-synonymous SNPs in exon 3	Autosomal overdominant	Kathrani et al. (2014)
Chronic superficial keratitis (CSK)	DLA-DRB1	Chromosome 12	Upstream regula-tory region SNP	Autosomal recessive	Barrientos et al. (2013)
Hemophilia A	Factor VIII	Chromosome X	Non-synonymous SNP in exon 11	X-linked	Christopherson et al. (2014)
Hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis (RCND)	BHD	Chromosome 5	Exon 7 change highly conserved amino acid	Autosomal dominant	Lingaas et al. (2003)

(continued)

Table 1 (continued)

Phenotype	Gene	Genomic location	Polymorphism	Mode of inheritance	Reference
Ichthyosis	ASPRV1	Chromosome 10	Non-synonymous SNP	De novo SNP (not inherited)	Bauer et al. (2017)
Inflammatory bowel disease	TLR4 and TLR5	Chromosome 11 (TLR4) and chromosome 38 (TLR5)	2 SNPs in TLR4 and 1 SNP in TLR5	TLR4 SNP recessive, TLR5 additive	Kathrani et al. (2010)
Olfactory ability	OR10H1-like and OR2K2-like	Chromosome 20 and chromosome 11	SNPs in exonic regions	Polygenic	Yang et al. (2016)
Pancreatic acinar atrophy	DLA-88	Chromosome 12	2 amino acid changes in protein	Autosomal recessive previously, now considered more complex	Tsai et al. (2013)
Pituitary dwarfism	LHX3	Chromosome 9	Deletion of a single 7 bp repeat in intron 5	Autosomal recessive	Voorbij et al. (2011)
White spotting	KIT	Chromosome 13	1 bp insertion in second exon	Autosomal dominant	Wong et al. (2013)

events. Prior to the advent of genomics and genome-wide association studies, artificial selection relied on phenotyping specific animals and breeding them for purposely bred traits. However, the combination of genotyping technology with genomic markers associated with phenotypes of interest allows genetically informed breeding plans to be developed to simultaneously maximize the phenotypes of interest while minimizing the time to achieve the desired artificial selection.

Many breed fanciers are actively working with kennel clubs and geneticists to try to breed out specific undesirable clinical phenotypes like cancer from their lines. This is a challenging process, and consequences of such approaches may result in unintended losses of heterozygosity and alleles within the breed. However, these consequences must be weighed against the backdrop of health for each breed. As the number of genetic markers implicated in dog traits continues to grow, the opportunities for breeding dogs with unique combinations of phenotypes will also increase. Novel breeds may emerge, that have a significantly reduced incidence of undesirable clinical phenotypes. Additionally, it is equally likely that designer dog breeds may be produced that possess unique combinations of morphological phenotypes that previously never co-occurred within the same breed. Combinatorial possibilities are quite literally endless.

Recently, designer dogs (hybrids of two different breeds) have come into fashion. Some dog fanciers view these emerging breeds as a destruction of the underlying breeds. However, others view these dogs as valuable companions and worthwhile pets. One example of such a designer dog is the Labradoodle, a dog produced by a cross of the Labrador Retriever with the Poodle. Considering the combinatorial explosion of pairs that can be crossed from 300 or 400 distinct dog breeds, there are between 44,850 and 79,800 distinct 2-breed designer dog breeds that can be produced from these 300 or 400 breeds, respectively. Furthermore, combining four different breeds to produce a hybrid dog results in more than one billion distinct four-breed combinations.

The demonstrated plasticity of the dog genome represents a powerful mechanism for creating and selecting phenotypes. It is likely that within another 1,000 years, dogs will be selected for combinations of phenotypes and traits that were once thought impossible. It will be truly exciting and breathtaking to witness the evolutionary journey humans will take with dogs.

Although to date dogs appear to have gone through two distinct selection processes, (1) an initial domestication followed by (2) an expansion of breeds more recently, beginning right now, dogs are entering the third selection process, one that will be carried out with the full scientific capability of the human species and where dogs end up will be anyone's guess.

The discoveries made in dog population genomics have been achieved using technology, such as genome sequencing, genotyping arrays, and gene expression arrays. This technology was developed in the past few decades. However, new genomics technology such as RNA sequencing, which provides advantages over microarray-based expression studies, will further open the window to understand complex patterns of gene expression associated with dog domestication, health, and disease. Additionally, the emerging tools associated with epigenetics will

Fig. 15 The human-animal bond was formed through the domestication of wolves into the companion animals we call dogs. Today, millions of dogs are members of human families. The strength of the human-animal bond is frequently represented in media, art, songs, movies, novels, paintings, sculptures, and family photos (as shown here)

undoubtedly provide a greater understanding of how phenotypic variation in dogs can arise through epigenetic regulation of genes. This information will elucidate the underlying mechanisms contributing to gene silencing and clarify why individuals with the same genotypes may exhibit strikingly different phenotypes.

In conclusion, the journey from speculation to knowledge has been very exciting. Moreover, although we have learned some new and important things about dogs, we still have much more to learn. Because dogs are considered to be the first species domesticated by humans, they are the ideal organism to study population genomics and unravel the mysteries underlying domestication and the impact artificial selection has had on anatomical, cognitive, dietary, social, behavioral and disease traits. Through thousands of years living among humans, dogs and humans have shared an extremely strong social bond (Fig. 15). The behavioral and cognitive basis for this bond is beginning to emerge from numerous studies aimed at deciphering the footprints of selection in the dog genome. This is a very exciting time for genomics and for dogs. As we gain a more detailed understanding of our interspecific relationship that evolved over the millennia, we will undoubtedly gain a scientific appreciation for what our hearts already know, and what we already know is that dogs are our best friends.

References

Ahonen SJ, Kaukonen M, Nussdorfer FD, Harman CD, Komaromy AM, Lohi H. A novel missense mutation in ADAMTS10 in Norwegian Elkhound primary glaucoma. PLoS One. 2014;9:e111941.

Appleyard SM, Hayward M, Young JI, Butler AA, Cone RD, Rubinstein M, Low MJ. A role for the endogenous opioid beta-endorphin in energy homeostasis. Endocrinology. 2003;144:1753–60.

Arendt M, Fall T, Lindblad-Toh K, Axelsson E. Amylase activity is associated with AMY2B copy numbers in dog: implications for dog domestication, diet and diabetes. Anim Genet. 2014;45:716–22.

Arendt ML, Melin M, Tonomura N, Koltookian M, Courtay-Cahen C, Flindall N, Bass J, Boerkamp K, Megquir K, Youell L, Murphy S, McCarthy C, London C, Rutteman GR, Starkey M, Lindblad-Toh K. Genome-wide association study of golden retrievers identifies germ-line risk factors predisposing to mast cell tumours. PLoS Genet. 2015;11:e1005647.

Arendt M, Cairns KM, Ballard JW, Savolainen P, Axelsson E. Diet adaptation in dog reflects spread of prehistoric agriculture. Heredity (Edinb). 2016;117:301–6.

Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar A, Lindblad-Toh K. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature. 2013;495:360–4.

Bannasch D, Young A, Myers J, Truvé K, Dickinson P, Gregg J, Davis R, Bongcam-Rudloff E, Webster MT, Lindblad-Toh K, Pedersen N. Localization of canine brachycephaly using an across breed mapping approach. PLoS One. 2010;5:e9632.

Barrientos LS, Zapata G, Crespi JA, Posik DM, Díaz S, It V, Peral-García P, Giovambattista G. A study of the association between chronic superficial keratitis and polymorphisms in the upstream regulatory regions of DLA-DRB1, DLA-DQB1 and DLA-DQA1. Vet Immunol Immunopathol. 2013;156:205–10.

Bauer A, Waluk DP, Galichet A, Timm K, Jagannathan V, Sayar BS, Wiener DJ, Dietschi E, Müller EJ, Roosje P, Welle MM, Leeb T. A de novo variant in the ASPRV1 gene in a dog with ichthyosis. PLoS Genet. 2017;13:e1006651.

Bence M, Marx P, Szantai E, Kubinyi E, Ronai Z, Banlaki Z. Lessons from the canine Oxtr gene: populations, variants and functional aspects. Genes Brain Behav. 2017;16:427–38.

Borge KS, Nord S, Van Loo P, Lingjærde OC, Gunnes G, Alnæs GI, Solvang HK, Lüders T, Kristensen VN, Børresen-Dale AL, Lingaas F. Canine mammary tumours are affected by

frequent copy number aberrations, including amplification of MYC and loss of PTEN. PLoS One. 2015;10:e0126371.

Boyko AR, et al. A simple genetic architecture underlies morphological variation in dogs. PLoS Biol. 2010;8:e1000451.

Cagan A, Blass T. Identification of genomic variants putatively targeted by selection during dog domestication. BMC Evol Biol. 2016;16:10.

Challis BG, Pritchard LE, Creemers JW, Delplanque J, Keogh JM, Luan J, Wareham NJ, Yeo GS, Bhattacharyya S, Froguel P, White A, Farooqi IS, O'Rahilly S. A missense mutation disrupting a dibasic prohormone processing site in pro-opiomelanocortin (POMC) increases susceptibility to early-onset obesity through a novel molecular mechanism. Hum Mol Genet. 2002;11:1997–2004.

Christopherson PW, Bacek LM, King KB, Boudreaux MK. Two novel missense mutations associated with hemophilia A in a family of Boxers, and a German Shepherd dog. Vet Clin Pathol. 2014;43:312–6.

Cooper AE, Ahonen S, Rowlan JS, Duncan A, Seppälä EH, Vanhapelto P, Lohi H, Komáromy AM. A novel form of progressive retinal atrophy in Swedish vallhund dogs. PLoS One. 2014;9:e106610.

Downs LM, Wallin-Håkansson B, Boursnell M, Marklund S, Hedhammar Å, Truvé K, Hübinette L, Lindblad-Toh K, Bergström T, Mellersh CS. A frameshift mutation in golden retriever dogs with progressive retinal atrophy endorses SLC4A3 as a candidate gene for human retinal degenerations. PLoS One. 2011;6:e21452.

Downs LM, Hitti R, Pregnolato S, Mellersh CS. Genetic screening for PRA-associated mutations in multiple dog breeds shows that PRA is heterogeneous within and between breeds. Vet Ophthalmol. 2014;17:126–30.

Drogemuller M, Jagannathan V, Dolf G, Butenhoff K, Kottmann-Berger S, Wess G, Leeb T. A single codon insertion in the PICALM gene is not associated with subvalvular aortic stenosis in Newfoundland dogs. Hum Genet. 2015;134:127–9.

Feldman R. The neurobiology of human attachments. Trends Cogn Sci. 2017;21:80–99.

Ferraresso S, et al. Epigenetic silencing of TFPI-2 in canine diffuse large B-cell lymphoma. PLoS One. 2014;9:e92707.

Freedman AH, et al. Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. 2014;10:e1004016.

Haas MA, Ngo L, Li SS, Schleich S, Qu Z, Vanyai HK, Cullen HD, Cardona-Alberich A, Gladwyn-Ng IE, Pagnamenta AT, Taylor JC, Stewart H, Kini U, Duncan KE, Teleman AA, Keays DA, Heng JI. De novo mutations in DENR disrupt neuronal development and link congenital neurological disorders to faulty mRNA translation re-initiation. Cell Rep. 2016;15:2251–65.

Hejjas K, Vas J, Topal J, Szantai E, Ronai Z, Szekely A, Kubinyi E, Horvath Z, Sasvari-Szekely M, Miklosi A. Association of polymorphisms in the dopamine D4 receptor gene and the activity-impulsivity endophenotype in dogs. Anim Genet. 2007;38:629–33.

Holder AL, Price JA, Adams JP, Volk HA, Catchpole B. A retrospective study of the prevalence of the canine degenerative myelopathy associated superoxide dismutase 1 mutation (SOD1:c.118G > A) in a referral population of German Shepherd dogs from the UK. Canine Genet Epidemiol. 2016;1:10.

Howrigan DP, Simonson MA, Keller MC. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. BMC Genomics. 2011;12:460.

Karyadi DM, Karlins E, Decker B, vonHoldt BM, Carpintero-Ramirez G, Parker HG, Wayne RK, Ostrander EA. A copy number variant at the KITLG locus likely confers risk for canine squamous cell carcinoma of the digit. PLoS Genet. 2013;9:e1003409.

Kathrani A, House A, Catchpole B, Murphy A, German A, Werling D, Allenspach K. Polymorphisms in the TLR4 and TLR5 gene are significantly associated with inflammatory bowel disease in German Shepherd dogs. PLoS One. 2010;5:e15740.

Kathrani A, Lee H, White C, Catchpole B, Murphy A, German A, Werling D, Allenspach K. Association between nucleotide oligomerisation domain two (Nod2) gene polymorphisms and canine inflammatory bowel disease. Vet Immunol Immunopathol. 2014;161:32–41.

Kirkness EF, Bafna V, Halpern AL, Levy S, Remington K, Rusch DB, Delcher AL, Pop M, Wang W, Fraser CM, Venter JC. The dog genome: survey sequencing and comparative analysis. Science. 2003;301:1898–903.

Kis A, Bence M, Lakatos G, Pergel E, Turcsán B, Pluijmakers J, Vas J, Elek Z, Brúder I, Földi L, Sasvári-Székely M, Miklósi A, Rónai Z, Kubinyi E. Oxytocin receptor gene polymorphisms are associated with human directed social behavior in dogs (Canis familiaris). PLoS One. 2014;9:e83993.

Kovacs K, Kis A, Pogany A, Koller D, Topal J. Differential effects of oxytocin on social sensitivity in two distinct breeds of dogs (Canis familiaris). Psychoneuroendocrinology. 2016;74:212–20.

Kubinyi E, Vas J, Hejjas K, Ronai Z, Brúder I, Turcsán B, Sasvari-Szekely M, Miklósi A. Polymorphism in the tyrosine hydroxylase (TH) gene is associated with activity-impulsivity in German Shepherd dogs. PLoS One. 2012;7:e30271.

Kuchtey J, Kunkel J, Esson D, Sapienza JS, Ward DA, Plummer CE, Gelatt KN, Kuchtey RW. Screening ADAMTS10 in dog populations supports Gly661Arg as the glaucoma-causing variant in beagles. Invest Ophthalmol Vis Sci. 2013;54:1881–6.

Lee YS, Challis BG, Thompson DA, Yeo GS, Keogh JM, Madonna ME, Wraight V, Sims M, Vatin V, Meyre D, Shield J, Burren C, Ibrahim Z, Cheetham T, Swift P, Blackwood A, Hung CC, Wareham NJ, Froguel P, Millhauser GL, O'Rahilly S, Farooqi IS. A POMC variant implicates beta-melanocyte-stimulating hormone in the control of human energy balance. Cell Metab. 2006;3:135–40.

Lee EJ, Merriwether DA, Kasparov AK, Nikolskiy PA, Sotnikova MV, Pavlova EY, Pitulko VV. Ancient DNA analysis of the oldest canid species from the Siberian Arctic and genetic contribution to the domestic dog. PLoS One. 2015;10:e0125759.

Li Y, Vonholdt BM, Reynolds A, Boyko AR, Wayne RK, Wu DD, Zhang YP. Artificial selection on brain-expressed genes during the domestication of dog. Mol Biol Evol. 2013;30:1867–76.

Li Y, Wang GD, Wang MS, Irwin DM, Wu DD, Zhang YP. Domestication of the dog from the wolf was promoted by enhanced excitatory synaptic plasticity: a hypothesis. Genome Biol Evol. 2014;6:3115–21.

Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ 3rd, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, PJ DJ, Kirkness E, Alvarez P, Biagi T, Brockman W, Butler J, Chin CW, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, Grabherr M, Kellis M, Kleber M, Bardeleben C, Goodstadt L, Heger A, Hitte C, Kim L, Koepfli KP, Parker HG, Pollinger JP, Searle SM, Sutter NB, Thomas R, Webber C, Baldwin J, Abebe A, Abouelleil A, Aftuck L, Ait-Zahra M, Aldredge T, Allen N, An P, Anderson S, Antoine C, Arachchi H, Aslam A, Ayotte L, Bachantsang P, Barry A, Bayul T, Benamara M, Berlin A, Bessette D, Blitshteyn B, Bloom T, Blye J, Boguslavskiy L, Bonnet C, Boukhgalter B, Brown A, Cahill P, Calixte N, Camarata J, Cheshatsang Y, Chu J, Citroen M, Collymore A, Cooke P, Dawoe T, Daza R, Decktor K, DeGray S, Dhargay N, Dooley K, Dooley K, Dorje P, Dorjee K, Dorris L, Duffey N, Dupes A, Egbiremolen O, Elong R, Falk J, Farina A, Faro S, Ferguson D, Ferreira P, Fisher S, FitzGerald M, Foley K, Foley C, Franke A, Friedrich D, Gage D, Garber M, Gearin G, Giannoukos G, Goode T, Goyette A, Graham J, Grandbois E, Gyaltsen K, Hafez N, Hagopian D, Hagos B, Hall J, Healy C, Hegarty R, Honan T, Horn A, Houde N, Hughes L, Hunnicutt L, Husby M, Jester B, Jones C, Kamat A, Kanga B, Kells C, Khazanovich D, Kieu AC, Kisner P, Kumar M, Lance K, Landers T, Lara M, Lee W, Leger JP, Lennon N, Leuper L, LeVine S, Liu J, Liu X, Lokyitsang Y, Lokyitsang T, Lui A, Macdonald J, Major J, Marabella R, Maru K, Matthews C, McDonough S, Mehta T, Meldrim J, Melnikov A, Meneus L, Mihalev A, Mihova T, Miller K, Mittelman R, Mlenga V, Mulrain L, Munson G, Navidi A, Naylor J, Nguyen T, Nguyen N, Nguyen C, Nguyen T, Nicol R, Norbu N, Norbu C, Novod N, Nyima T, Olandt P, O'Neill B, O'Neill K, Osman S, Oyono L, Patti C, Perrin D,

Phunkhang P, Pierre F, Priest M, Rachupka A, Raghuraman S, Rameau R, Ray V, Raymond C, Rege F, Rise C, Rogers J, Rogov P, Sahalie J, Settipalli S, Sharpe T, Shea T, Sheehan M, Sherpa N, Shi J, Shih D, Sloan J, Smith C, Sparrow T, Stalker J, Stange-Thomann N, Stavropoulos S, Stone C, Stone S, Sykes S, Tchuinga P, Tenzing P, Tesfaye S, Thoulutsang D, Thoulutsang Y, Topham K, Topping I, Tsamla T, Vassiliev H, Venkataraman V, Vo A, Wangchuk T, Wangdi T, Weiand M, Wilkinson J, Wilson A, Yadav S, Yang S, Yang X, Young G, Yu Q, Zainoun J, Zembek L, Zimmer A, Lander ES. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature. 2005;438:803–19.

Lingaas F, Comstock KE, Kirkness EF, Sørensen A, Aarskaug T, Hitte C, Nickerson ML, Moe L, Schmidt LS, Thomas R, Breen M, Galibert F, Zbar B, Ostrander EA. A mutation in the canine BHD gene is associated with hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis in the German Shepherd dog. Hum Mol Genet. 2003;12:3043–53.

Massey J, Short AD, Catchpole B, House A, Day MJ, Lohi H, Ollier WE, Kennedy LJ. Genetics of canine anal furunculosis in the German Shepherd dog. Immunogenetics. 2014;66:311–24.

Melin M, Rivera P, Arendt M, Elvers I, Murén E, Gustafson U, Starkey M, Borge KS, Lingaas F, Häggström J, Saellström S, Rönnberg H, Lindblad-Toh K. Genome-wide analysis identifies germ-line risk factors associated with canine mammary tumours. PLoS Genet. 2016;12:e1006029.

Meurs KM, Lahmers S, Keene BW, White SN, Oyama MA, Mauceli E, Lindblad-Toh K. 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. Hum Genet. 2012;131:1319–25.

Monteagudo LV, Tejedor MT. The b(c) allele of TYRP1 is causative for the recessive brown (liver) colour in German Shepherd dogs. Anim Genet. 2015;46:588–9.

Mosher DS, Quignon P, Bustamante CD, Sutter NB, Mellersh CS, Parker HG, Ostrander EA. A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. PLoS Genet. 2007;3:e79.

Nagasawa M, Mitsui S, En S, Ohtani N, Ohta M, Sakuma Y, Onaka T, Mogi K, Kikusui T. Social evolution. Oxytocin-gaze positive loop and the coevolution of human-dog bonds. Science. 2015;348:333–6.

Oliver JA, Forman OP, Pettitt L, Mellersh CS. Two independent mutations in ADAMTS17 are associated with primary open angle glaucoma in the Basset Hound and Basset Fauve de Bretagne breeds of dog. PLoS One. 2015;10:e0140436.

Ovodov ND, Crockford SJ, Kuzmin YV, Higham TF, Hodgins GW, van der Plicht J. A 33,000-year-old incipient dog from the Altai Mountains of Siberia: evidence of the earliest domestication disrupted by the last glacial maximum. PLoS One. 2011;6:e22821.

Parker HG, Ostrander EA. Canine genomics and genetics: running with the pack. PLoS Genet. 2005;1:e58.

Persson ME, Wright D, Roth LS, Batakis P, Jensen P. Genomic regions associated with interspecies communication in dogs contain genes related to human social disorders. Sci Rep. 2016;6:33439.

Philipp U, Vollmar A, Haggstrom J, Thomas A, Distl O. Multiple Loci are associated with dilated cardiomyopathy in Irish wolfhounds. PLoS One. 2012;7:e36691.

Quignon P, Herbin L, Cadieu E, Kirkness EF, Hédan B, Mosher DS, Galibert F, André C, Ostrander EA, Hitte C. Canine population structure: assessment and impact of intra-breed stratification on SNP-based association studies. PLoS One. 2007;2:e1324.

Raffan E, Dennis RJ, O'Donovan CJ, Becker JM, Scott RA, Smith SP, Withers DJ, Wood CJ, Conci E, Clements DN, Summers KM, German AJ, Mellersh CS, Arendt ML, Iyemere VP, Withers E, Söder J, Wernersson S, Andersson G, Lindblad-Toh K, Yeo GS, O'Rahilly S. A deletion in the canine POMC gene is associated with weight and appetite in obesity-prone labrador retriever dogs. Cell Metab. 2016;23:893–900.

Reiter T, Jagoda E, Capellini TD. Dietary variation and evolution of gene copy number among dog breeds. PLoS One. 2016;11:e0148899.

Romero T, Nagasawa M, Mogi K, Hasegawa T, Kikusui T. Oxytocin promotes social bonding in dogs. Proc Natl Acad Sci U S A. 2014;111:9085–90.

Saetre P, Lindberg J, Leonard JA, Olsson K, Pettersson U, Ellegren H, Bergström TF, Vilà C, Jazin E. From wild wolf to domestic dog: gene expression changes in the brain. Brain Res Mol Brain Res. 2004;126:198–206.

Schoenebeck JJ, Hutchinson SA, Byers A, Beale HC, Carrington B, Faden DL, Rimbault M, Decker B, Kidd JM, Sood R, Boyko AR, Fondon JW 3rd, Wayne RK, Bustamante CD, Ciruna B, Ostrander EA. Variation of BMP3 contributes to dog breed skull diversity. PLoS Genet. 2012;8:e1002849.

Shannon LM, Boyko RH, Castelhano M, Corey E, Hayward JJ, McLean C, White ME, Abi Said M, Anita BA, Bondjengo NI, Calero J, Galov A, Hedimbi M, Imam B, Khalap R, Lally D, Masta A, Oliveira KC, Pérez L, Randall J, Tam NM, Trujillo-Cornejo FJ, Valeriano C, Sutter NB, Todhunter RJ, Bustamante CD, Boyko AR. Genetic structure in village dogs reveals a Central Asian domestication origin. Proc Natl Acad Sci U S A. 2015;112:13639–44.

Shearin AL, Ostrander EA. Canine morphology: hunting for genes and tracking mutations. PLoS Biol. 2010;8:e1000310.

Stern JA, White SN, Lehmkuhl LB, Reina-Doreste Y, Ferguson JL, Nascone-Yoder NM, Meurs KM. A single codon insertion in PICALM is associated with development of familial subvalvular aortic stenosis in Newfoundland dogs. Hum Genet. 2014;133:1139–48.

Stern JA, Hsue W, Song KH, Ontiveros ES, Luis Fuentes V, Stepien RL. Severity of mitral valve degeneration is associated with chromosome 15 loci in whippet dogs. PLoS One. 2015;10:e0141234.

Tengvall K, Kierczak M, Bergvall K, Olsson M, Frankowiack M, Farias FH, Pielberg G, Carlborg Ö, Leeb T, Andersson G, Hammarström L, Hedhammar Å, Lindblad-Toh K. Genome-wide analysis in German Shepherd dogs reveals association of a locus on CFA 27 with atopic dermatitis. PLoS Genet. 2013;9:e1003475.

Tengvall K, Kozyrev S, Kierczak M, Bergvall K, Farias FH, Ardesjö-Lundgren B, Olsson M, Murén E, Hagman R, Leeb T, Pielberg G, Hedhammar Å, Andersson G, Lindblad-Toh K. Multiple regulatory variants located in cell type-specific enhancers within the PKP2 locus form major risk and protective haplotypes for canine atopic dermatitis in German Shepherd dogs. BMC Genet. 2016;17:97.

Thalmann O, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL, Germonpré MB, Sablin MV, López-Giráldez F, Domingo-Roura X, Napierala H, Uerpmann HP, Loponte DM, Acosta AA, Giemsch L, Schmitz RW, Worthington B, Buikstra JE, Druzhkova A, Graphodatsky AS, Ovodov ND, Wahlberg N, Freedman AH, Schweizer RM, Koepfli KP, Leonard JA, Meyer M, Krause J, Pääbo S, Green RE, Wayne RK. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. Science. 2013;342:871–4.

Toro R, Pérez-Serra A, Campuzano O, Moncayo-Arlandi J, Allegue C, Iglesias A, Mangas A, Brugada R. Familial dilated cardiomyopathy caused by a novel frameshift in the BAG3 gene. PLoS One. 2016;11:e0158730.

Truvé K, Dickinson P, Xiong A, York D, Jayashankar K, Pielberg G, Koltookian M, Murén E, Fuxelius HH, Weishaupt H, Swartling FJ, Andersson G, Hedhammar Å, Bongcam-Rudloff E, Forsberg-Nilsson K, Bannasch D, Lindblad-Toh K. Utilizing the dog genome in the search for novel candidate genes involved in glioma development-genome wide association mapping followed by targeted massive parallel sequencing identifies a strongly associated locus. PLoS Genet. 2016;12:e1006000.

Tsai KL, Starr-Moss AN, Venkataraman GM, Robinson C, Kennedy LJ, Steiner JM, Clark LA. Alleles of the major histocompatibility complex play a role in the pathogenesis of pancreatic acinar atrophy in dogs. Immunogenetics. 2013;65:501–9.

Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Rosengren Pielberg G, Sigurdsson S, Fall T, Seppälä EH, Hansen MS, Lawley CT, Karlsson EK, Consortium LUPA, Bannasch D, Vilà C, Lohi H, Galibert F, Fredholm M, Häggström J, Hedhammar A, André C, Lindblad-Toh K,

Hitte C, Webster MT. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. PLoS Genet. 2011;7:e1002316.

Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, Reynolds A, Bryc K, Brisbin A, Knowles JC, Mosher DS, Spady TC, Elkahloun A, Geffen E, Pilot M, Jedrzejewski W, Greco C, Randi E, Bannasch D, Wilton A, Shearman J, Musiani M, Cargill M, Jones PG, Qian Z, Huang W, Ding ZL, Zhang YP, Bustamante CD, Ostrander EA, Novembre J, Wayne RK. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature. 2010;464:898–902.

Voorbij AM, van Steenbeek FG, Vos-Loohuis M, Martens EE, Hanson-Nilsson JM, van Oost BA, Kooistra HS, Leegwater PA. A contracted DNA repeat in LHX3 intron 5 is associated with aberrant splicing and pituitary dwarfism in German Shepherd dogs. PLoS One. 2011;6:e27940.

Wang GD, Zhai W, Yang HC, Fan RX, Cao X, Zhong L, Wang L, Liu F, Wu H, Cheng LG, Poyarkov AD, Poyarkov NA Jr, Tang SS, Zhao WM, Gao Y, Lv XM, Irwin DM, Savolainen P, Wu CI, Zhang YP. The genomics of selection in dogs and the parallel evolution between dogs and humans. Nat Commun. 2013;4:1860.

Wiik AC, Ropstad EO, Eksten B, Karlstam L, Wade CM, Lingaas F. Progressive retinal atrophy in Shetland sheepdog is associated with a mutation in the CNGA1 gene. Anim Genet. 2015;46:515–21.

Wijesena HR, Schmutz SM. A missense mutation in SLC45A2 is associated with albinism in several small long haired dog breeds. J Hered. 2015;106:285–8.

Winkler PA, Gornik KR, Ramsey DT, Dubielzig RR, Venta PJ, Petersen-Jones SM, Bartoe JT. A partial gene deletion of SLC45A2 causes oculocutaneous albinism in Doberman pinscher dogs. PLoS One. 2014;9:e92127.

Wolf ZT, Brand HA, Shaffer JR, Leslie EJ, Arzi B, Willet CE, Cox TC, McHenry T, Narayan N, Feingold E, Wang X, Sliskovic S, Karmi N, Safra N, Sanchez C, Deleyiannis FW, Murray JC, Wade CM, Marazita ML, Bannasch DL. Genome-wide association studies in dogs and humans identify ADAMTS20 as a risk variant for cleft lip and palate. PLoS Genet. 2015;11:e1005059.

Wong AK, Ruhe AL, Robertson KR, Loew ER, Williams DC, Neff MW. A de novo mutation in KIT causes white spotting in a subpopulation of German Shepherd dogs. Anim Genet. 2013;44:305–10.

Yang M, Geng GJ, Zhang W, Cui L, Zhang HX, Zheng JL. SNP genotypes of olfactory receptor genes associated with olfactory ability in German Shepherd dogs. Anim Genet. 2016;47:240–4.