

# Commonalities in Development of Pure Breeds and Population Isolates Revealed in the Genome of the Sardinian Fonnì's Dog

Dayna L. Dreger,\* Brian W. Davis,\* Raffaella Cocco,<sup>†</sup> Sara Sechi,<sup>†</sup> Alessandro Di Cerbo,<sup>†</sup> Heidi G. Parker,\* Michele Polli,<sup>§</sup> Stefano P. Marelli,<sup>§</sup> Paola Crepaldi,<sup>§</sup> and Elaine A. Ostrander\*<sup>1</sup>

\*Cancer Genetics and Comparative Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892; <sup>†</sup>Dipartimento di Medicina Veterinaria, Sezione Clinica Medica, Università degli Studi di Sassari, 07100, Italy; <sup>‡</sup>Dipartimento di Scienze Mediche, Orali e Biotecnologiche, Specializzazione Biochimica Clinica, Università degli Studi G. d'Annunzio Chieti Pescara, 66100, Italy; and <sup>§</sup>Dipartimento di Medicina Veterinaria, Università degli Studi di Milano, 20133, Italy

**ABSTRACT** The island inhabitants of Sardinia have long been a focus for studies of complex human traits due to their unique ancestral background and population isolation reflecting geographic and cultural restriction. Population isolates share decreased genomic diversity, increased linkage disequilibrium, and increased inbreeding coefficients. In many regions, dogs and humans have been exposed to the same natural and artificial forces of environment, growth, and migration. Distinct dog breeds have arisen through human-driven selection of characteristics to meet an ideal standard of appearance and function. The Fonnì's Dog, an endemic dog population on Sardinia, has not been subjected to an intensive system of artificial selection, but rather has developed alongside the human population of Sardinia, influenced by geographic isolation and unregulated selection based on its environmental adaptation and aptitude for owner-desired behaviors. Through analysis of 28 dog breeds, represented with whole-genome sequences from 13 dogs and ~170,000 genome-wide single nucleotide variants from 155 dogs, we have produced a genomic illustration of the Fonnì's Dog. Genomic patterns confirm within-breed similarity, while population and demographic analyses provide spatial identity of Fonnì's Dog to other Mediterranean breeds. Investigation of admixture and fixation indices reveals insights into the involvement of Fonnì's Dogs in breed development throughout the Mediterranean. We describe how characteristics of population isolates are reflected in dog breeds that have undergone artificial selection, and are mirrored in the Fonnì's Dog through traditional isolating factors that affect human populations. Lastly, we show that the genetic history of Fonnì's Dog parallels demographic events in local human populations.

**KEYWORDS** dog; whole-genome sequence; demography; population structure

**T**HE domestic dog has undergone intensive human-driven selective pressure in an effort to develop canines of a particular appearance or with a desired behavior pattern (Galibert *et al.* 2011; Wayne and vonHoldt 2012). While considerable effort has gone into decoding the genetic basis of morphologic traits that vary within dog populations, such as body size or coat color, the challenge of identifying genomic features that distinguish breeds selected to support human survival is largely unmet (summarized in Boyko 2011;

Rimbault and Ostrander 2012; Schoenebeck and Ostrander 2014). Breeds specifically suited to perform such tasks include those that excel at hunting, herding, and protection of property. An appeal to aesthetics and functional design has also resulted in selection of dogs with specific and uniform morphology. Since much of modern breed formation took place 100–300 years ago (Fogel 1995), restrictive geography and burgeoning industry have played a part in influencing the distribution of genetic variation among isolated populations.

From a genetic viewpoint, dog breeds are defined as populations with human-induced or geographic barriers to gene flow, and fixed traits that will reliably reproduce selected characteristics through multiple generations. There are several organizations that recognize and register dog breeds. One of the largest, both in terms of breed recognition and geographic reach, is the Federation Cynologique Internationale (FCI), an

Copyright © 2016 by the Genetics Society of America  
doi: 10.1534/genetics.116.192427

Manuscript received June 8, 2016; accepted for publication August 10, 2016;  
published Early Online August 12, 2016.

Available freely online through the author-supported open access option.

<sup>1</sup>Corresponding author: National Human Genome Research Institute, National Institutes of Health, Bldg. 50, Room 5351, 50 South Dr., Bethesda, MD 20892. E-mail: eostrand@mail.nih.gov

international organization encompassing 91 member countries and recognizing 343 breeds (<http://www.fci.be/en/Presentation-of-our-organisation-4.html>). The FCI, however, does not necessarily recognize regional varieties of dogs that exist worldwide, which are often acknowledged at only a local or niche level. Many of these dog populations have often undergone selection largely for functional purposes and behaviors that are distinctly important to the humans in the region and, as such, they represent isolated and unique genetic histories.

The study of human island populations has been of considerable interest to geneticists, as such populations often exhibit decreased genomic diversity (Sajantila *et al.* 1996; Capocasa *et al.* 2014; Di Gaetano *et al.* 2014), an increase in linkage disequilibrium (LD) (Bendjilali *et al.* 2014), and increased inbreeding coefficients (Zhai *et al.* 2016). These same population metrics have proven informative for the study of complex traits and heritable diseases, particularly in the context of relatively young populations that have experienced rapid growth, share a common environment and cultural practices, and frequently maintain detailed genealogical records (reviewed in Peltonen *et al.* 2000). Populations can become genetically isolated through reproductive restraints imposed by cultural or religious practices, language, or industry. In these instances, societal restraints only influence the human gene pool, rather than other cohabitant organisms. Geographic isolation, however, is an additional barrier that can affect both the human inhabitants of a region, as well as the endemic or introduced flora and fauna. The Italian island of Sardinia experiences both isolating mechanisms, providing a unique opportunity to investigate complex traits in a human population with low genetic diversity (Capocasa *et al.* 2014; Di Gaetano *et al.* 2014), high frequency of private variants (Sidore *et al.* 2015), and unique genomic ancestry (Fiorito *et al.* 2015).

The country of Italy is the site of origin for several internationally recognized modern breeds such as the Cane Corso, Maltese, Neapolitan Mastiff, Spinone Italiano, Bracco Italiano, and Cirneco dell'Etna. Italy also stakes claim to a great number of regional varieties, including the Fonní's Dog and the Mastino Abruzzese, both livestock guardians; the Cane Paratore, a herding breed; and the Levriero Meridionale, a sighthound. These breeds primarily exist as isolates surrounding their historic regions of origin, not having gained popularity outside local breed enthusiasts.

Among the most unique of the Italian varieties is the Fonní's Dog, also known as the Cane Fonnese, Pastore Fonnese, or Sardinian Shepherd Dog. It is a large livestock and property guardian breed originating from the region surrounding the city of Fonní. Present in rough- and smooth-coated varieties and in varying colors, the unifying features of the breed include a characteristically intense facial expression and instinctive propensity toward guarding behaviors and wariness of strangers. Historical accounts portray dogs fitting this description residing in Fonní and the surrounding regions in the mid- to late-19th century (Tyndale 1849; Bresciani 1850; Cetti 1774; Edwardes 1889).

International breed recognition of the Fonní's Dog is being pursued by a dedicated group of breeders and enthusiasts

(<http://www.canefonnese.it/>), with a goal of preserving the distinct heritage of this remarkable breed. Preliminary studies (Sechi *et al.* 2016) have characterized the morphological commonalities of the Fonní's Dog, showing that they are consistent with features of a true-breeding population. Studies using a limited number of microsatellite markers support this hypothesis (Sechi *et al.* 2016).

In this study, we have used whole-genome sequence (WGS) at an average coverage of 42× over 13 canids, together with SNP analysis of 173,662 variants across 155 canids, to characterize the genomic architecture of Fonní's Dogs in the context of geographically and historically proximate dog breeds (Figure 1). We investigate the similarities between human population isolates and dog breeds and, subsequently, breed development when driven by human-defined regulation of selection or when predominantly based on behavioral aptitude for guardian functions in the absence of aesthetic preference. We demonstrate the propensity for Fonní's Dogs to display genomic characteristics equivalent to those of established and acknowledged breeds, providing a dynamic account of how geographic isolation and behavior-driven selection function to produce unique breed populations. Finally, we reveal parallelisms between the breed foundation of the Fonní's Dog and the human population demographics of Sardinia, exposing an adroit consideration for investigations of population structure and the breadth of research application for dog breeds as population isolates.

## Materials and Methods

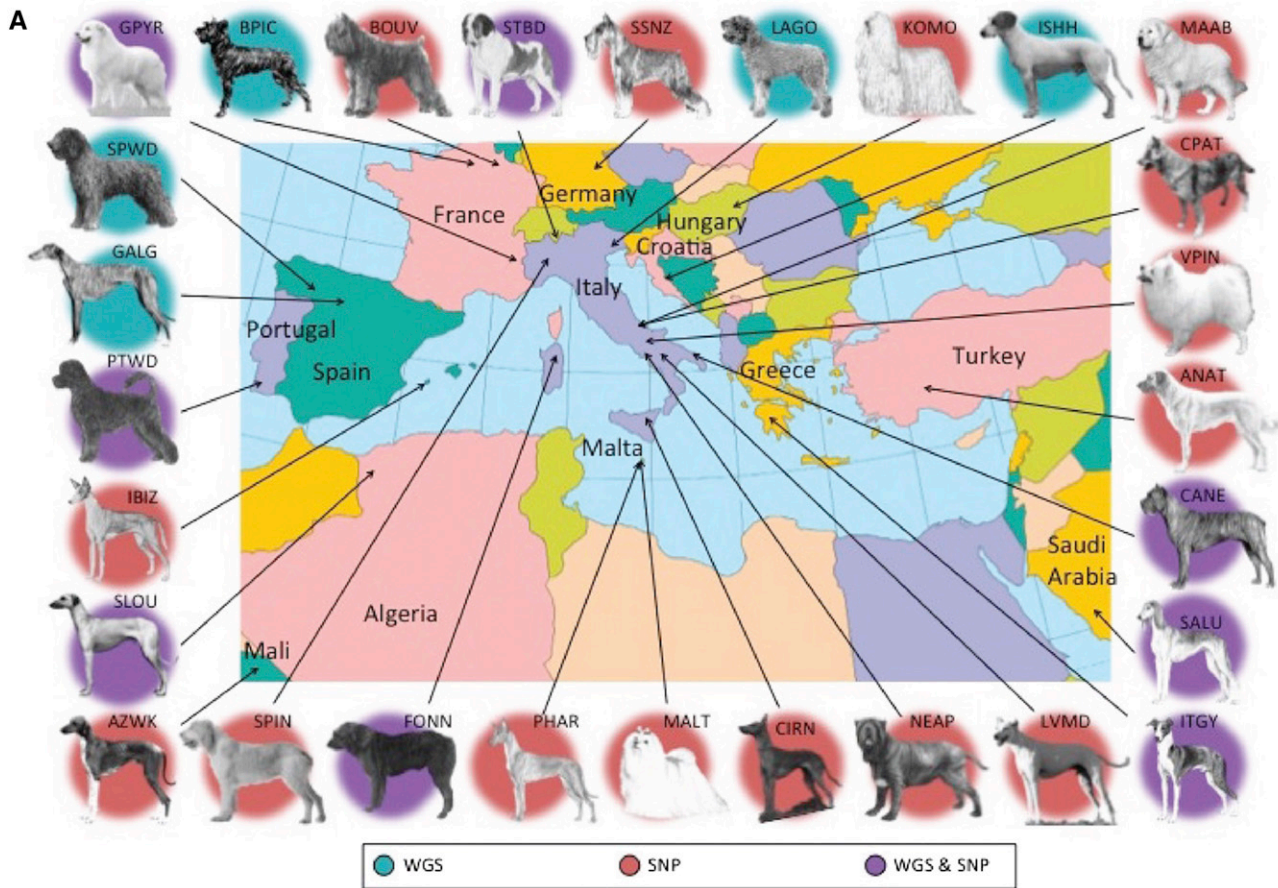
### Sample acquisition and genotyping

Blood samples were obtained and sent to the National Human Genome Research Institute (NHGRI) from six unrelated Fonní's Dogs, one each from the Sardinian regions of Fonní, Ozieri, Cagliari, Sassari, Porto Torres, and Nuoro (Figure 1). Remaining dog breed samples were submitted by owners with a signed consent, in accordance with the NHGRI Animal Care and Use Committee. Genomic DNA was extracted from all blood samples using a standard phenol chloroform method, aliquoted, and stored at  $-80^{\circ}$  (Sambrook *et al.* 1989). Extractions of DNA from saliva collections were performed using the recommended Performagene protocol (DNA Genotek, Ottawa, ON).

A set of 135 dogs representing 19 breeds (Table 1) were genotyped using the Illumina (San Diego, CA) CanineHD SNP chip, which has 173,662 potentially informative markers. Genotype calling was conducted in Illumina Genome Studio, with a protocol specifying a  $>90\%$  call rate and GenTrain score of  $>0.4$ . Illumina Canine HD SNP genotypes for eight Maltese, two Komondors, two Spinone Italiano, and eight Bouvier des Flandres were obtained from a previous publication (Hayward *et al.* 2016). We thus analyzed data from 155 SNP-genotyped dogs representing 23 breeds.

### WGS alignment

Publicly available WGS data were obtained from the Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>) for



**Figure 1** Dog breeds used in molecular analyses. (A) Geographic representation of the region of origin for the 23 Mediterranean breeds. The (B) smooth-coated and (C) rough-coated varieties are both recognized in Fanni's Dogs. Breed abbreviations are listed in Table 1.

a total of 12 dogs representing 12 breeds (Table 1). WGS from one Fanni's Dog was produced for this study by the National Institutes of Health Intramural Sequencing Center

using the Illumina TruSeq DNA PCR-Free Protocol (Cat. #FC-121-3001). Data were aligned to the CanFam 3.1 reference genome (<http://genome.ucsc.edu/cgi-bin/hgGateway?db=canFam3>)

**Table 1 Type and number of dogs used in analyses**

Breed	Abbreviation	Process <sup>a</sup>	No.	WGS source
Anatolian Shepherd	ANAT	SNP	6	
Azawakh Hound	AZWK	SNP	5	
Berger Picard	BPIC	WGS	1	SRR2016171
Bouvier des Flandres	BOUV	SNP	8	
Cane Corso	CANE	WGS	1	SRR2747522
		SNP	9	
Cane Paratore	CPAT	SNP	2	
Cirneco dell'Etna	CIRN	SNP	5	
Fonni's Dog	FONN	WGS	1	PRJNA318762
		SNP	6	
Great Pyrenees	GPYR	WGS	1	SAMN03801670
		SNP	10	
Ibizan Hound	IBIZ	SNP	10	
Istrian Shorthaired Hound	ISHH	WGS	1	SAMN02485584
Italian Greyhound	ITGY	WGS	1	SAMN03801673
		SNP	10	
Komondor	KOMO	SNP	2	
Labrador Retriever	LAGO	WGS	1	PRJNA318762
Levriero Meridionale	LVMD	SNP	2	
Maltese	MALT	SNP	10	
Mastino Abruzzese	MAAB	SNP	2	
Neapolitan Mastiff	NEAP	SNP	6	
Pharaoh Hound	PHAR	SNP	2	
Portuguese Water Dog	PTWD	WGS	1	PRJNA318762
		SNP	10	
Saluki	SALU	WGS	1	SAMN03801686
		SNP	19	
Sloughi	SLOU	WGS	1	PRJNA318762
		SNP	5	
Spanish Galgo	GALG	WGS	1	SAMN03168380
Spanish Water Dog	SPWD	WGS	1	PRJEB7903
Spinone Italiano	SPIN	SNP	2	
Standard Schnauzer	SSNZ	SNP	10	
Saint Bernard	STBD	WGS	1	PRJNA263947
		SNP	10	
Volpino Italiano	VPIN	SNP	4	

SRA accession numbers are included for all WGS.

<sup>a</sup> This column indicates individuals genotyped on the Illumina HD Canine SNP array (SNP) or WGS.

using Burrows-Wheeler Aligner 0.7.10 MEM (Li and Durbin 2009). SAMtools 0.1.10 (Li *et al.* 2009) was used for sorting, and PicardTools 1.119 (<https://github.com/broadinstitute/picard>) for screening for putative PCR duplicate reads. Local realignment was conducted with Genome Analysis Toolkit (GATK) 3.2-2 (DePristo *et al.* 2011) and based on documented and novel indels (Axelsson *et al.* 2013). Training sets of dbSNP and Illumina CanineHD chip positions were used for base quality recalibration. Single nucleotide variants were called with gVCF mode of HaplotypeCaller (Van der Auwera *et al.* 2013) for each individual dog, and again jointly across all sequenced dogs. Variant quality score recalibration was conducted with GATK best practices and default parameters, and the initial alignment training sets. Resultant jointly called variant call formats (VCFs) were filtered for CpG islands, gaps, and repeats, based on CanFam 3.1 reference genome annotations (<http://genome.ucsc.edu/cgi-bin/hgGateway?db=canFam3>).

The previously unpublished WGS from the Fonni's Dog has been uploaded to the Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>). Variants were annotated with SNPeff (Cingolani *et al.* 2012) based on CanFam 3.1.76 (Lindblad-Toh *et al.* 2005).

### Genomic characterization of breeds

Inbreeding coefficients and homozygosity were calculated from the SNP data using the “het” and “homozyg” functions of PLINK v.1.07 (Purcell *et al.* 2007), respectively. Genome-wide length of homozygosity was measured using sliding windows of 5 Mb, allowing for one heterozygous call and a maximum of five missing genotypes per window. Individual regions of homozygosity were identified from SNP genotypes using PLINK v.1.07 software and the homozyg function, and required a minimum of 10 sequential SNPs with a heterozygous allowance of 1. Haplotypes were analyzed from regions that were shared by four of the six Fonni's Dogs, and smaller homozygous regions of three or more sequential SNPs that were shared by all six Fonni's Dogs were identified.

Principle components analysis (PCA) was calculated with Eigensoft v.6.0.1 (Patterson *et al.* 2006; Price *et al.* 2006) and PCA plots were drawn using Partek Genomics Suite v.6.6 (Partek, St. Louis, MO). FastSTRUCTURE v.1.0 (Raj *et al.* 2014) was used to identify population structure from SNP data. Neighbor-joining phylogeny was constructed with Phylip v.3.696 (Felsenstein 1989) using a distance matrix from PLINK v.1.07 (Purcell *et al.* 2007) with 100 bootstrapped repetitions, and dendrograms of the consensus tree were drawn in FigTree v.1.4.2 (Rambaut 2014). VCFtools v.0.1.14 (Danecek *et al.* 2011) was used to calculate Weir and Cockerham's  $F_{ST}$  for every breed-to-breed pair. Effective population size ( $N_e$ ) of each breed with greater than four dogs was estimated through SNP-based LD analysis with SNeP (Barbato *et al.* 2015). Private variants were called for each individual dog, relative to the pool of 13 WGSs representing 13 breeds, using the “singletons” filter of VCFtools v.0.1.14 (Danecek *et al.* 2011). These variants will be referred to as “individual” variants throughout, to distinguish them as unique to single dogs, though not necessarily indicative of breed-specific private variants.

Admixture for every breed-to-breed combination was assessed using the “three\_pop” function of TreeMix v.1.12 (Pickrell and Pritchard 2012; Pickrell *et al.* 2012) and AdmixTools v.3.0 (Alexander *et al.* 2009) using the SNP data, and Analysis of Next Generation Sequencing Data (ANGSD) v.0.911 (Korneliussen *et al.* 2014) using the WGS data. Z scores obtained from Patterson's  $D$  and  $f_3$  for introgressions involving Fonni's Dog were compared across computational algorithms, and admixing events suggested by two or more programs were identified for further consideration. A critical Z value of  $|Z| = 3$  was used to determine significance. The SNP data from the Mediterranean breeds was analyzed in TreeMix v.1.12 to produce phylogenies with the most likely introgression events, allowing for 1 through 10, 15, and 25 possible migrations.

## Data availability

WGS was produced for this study from a single Fonní's Dog, submitted to the Sequence Read Archive (PRJNA318762) for release upon manuscript acceptance. Previously published WGS obtained from the Sequence Read Archive is listed in Table 1 with corresponding accession numbers. Genotype data from the Illumina CanineHD array has been submitted to Gene Expression Omnibus (GSE83160).

## Results

### *A representative population of dog breeds originating in the Mediterranean*

A selection of 28 dog breeds, originating from regions surrounding the Mediterranean Sea, was chosen for genomic population analysis (Figure 1). Numbers of dogs per breed ranged from 1 (Istrian Shorthaired Hound, Lagotto Romagnolo, Spanish Water Dog, Spanish Galgo) to 19 (Saluki). The Fonní's Dog, Cane Paratore, Mastino Abruzzese, Volpino Italiano, Levriero Meridionale, and Cirneco dell'Etna were sampled from Italy; the Sloughi were collected from North Africa; and the Azawakh from Mali. The remaining breeds were predominantly collected from the United States and are largely American Kennel Club-recognized breeds. The entirety of these breeds is referred to herein as the "Mediterranean group," while a subset, consisting of those breeds with written or assumed history tracing to Italy (Cane Corso, Cane Paratore, Cirneco dell'Etna, Fonní's Dog, Levriero Meridionale, Maltese, Mastino Abruzzese, Neapolitan Mastiff, Pharaoh Hound, Spinone Italiano, and Volpino Italiano) (<http://www.enci.it/libro-genealogico/razze-italiane>; <http://www.bordercolliemuseum.org/BCCousins/EuropeWestern/Italy.html>), are also classified specifically as the "Italian group." While tracing its origin to Hungary, previous analyses (unpublished data) have repeatedly shown shared genetic identity between the Komondor breed and those in the Italian group. For the purposes of this study, the Komondor has been included in the Italian group. All of the above are termed "breeds" for the purposes of this study. Note, however, that a subset of breeds (Mastino Abruzzese, Cane Paratore, and Levriero Meridionale) is not officially recognized by any formal registry, while the remainder of the breeds have received national or international recognition.

### *SNP-based measures of homozygosity for Mediterranean breeds*

Since decreased levels of genetic diversity can be indicative of selective breeding, founder effects, or population bottlenecks, and therefore breed formation, molecular analyses included measures of homozygosity and inbreeding coefficients for the Mediterranean breeds. The number of homozygous regions from the Mediterranean breeds (Figure 2, Table 2) with a minimum length of 1 Mb, ranged from a breed mean of 12 (Mastino Abruzzese) to 114 (Saint Bernard), with a mean of 67 across all breeds. The total length of homozygosity,

calculated as the sum of the lengths for each of the previously determined regions of homozygosity per dog and averaged across each breed, ranged from a breed mean of 101.8 Mb (Mastino Abruzzese) to 747.1 Mb (Pharaoh Hound), with a mean of 466.9 Mb across all dogs. The Fonní's Dog had a breed range of 8–49 homozygous regions and 59.6–760.6 Mb total homozygosity, with within-breed means of 26 and 348.7 Mb, respectively (Table 2).

Individual inbreeding coefficients had a maximum within-breed mean of 0.39 (Pharaoh Hound) and a minimum of 0.03 (Mastino Abruzzese), with a mean of 0.22 across all breeds (Figure 2, Table 2). The Fonní's Dog ranged from 0.01 to 0.30, and had a breed mean inbreeding coefficient of 0.13. Thus, the Fonní's Dog presented homozygosity and inbreeding values within the range of equivalent values expressed by other Mediterranean dog breeds.

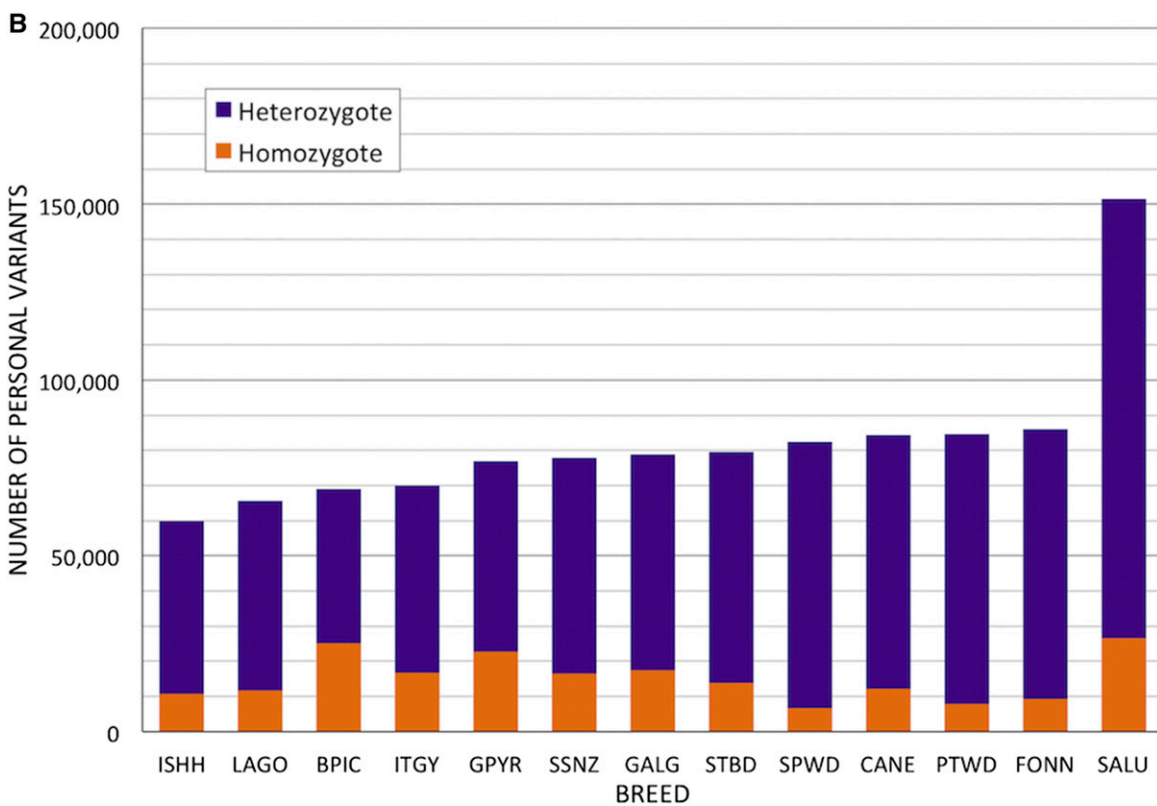
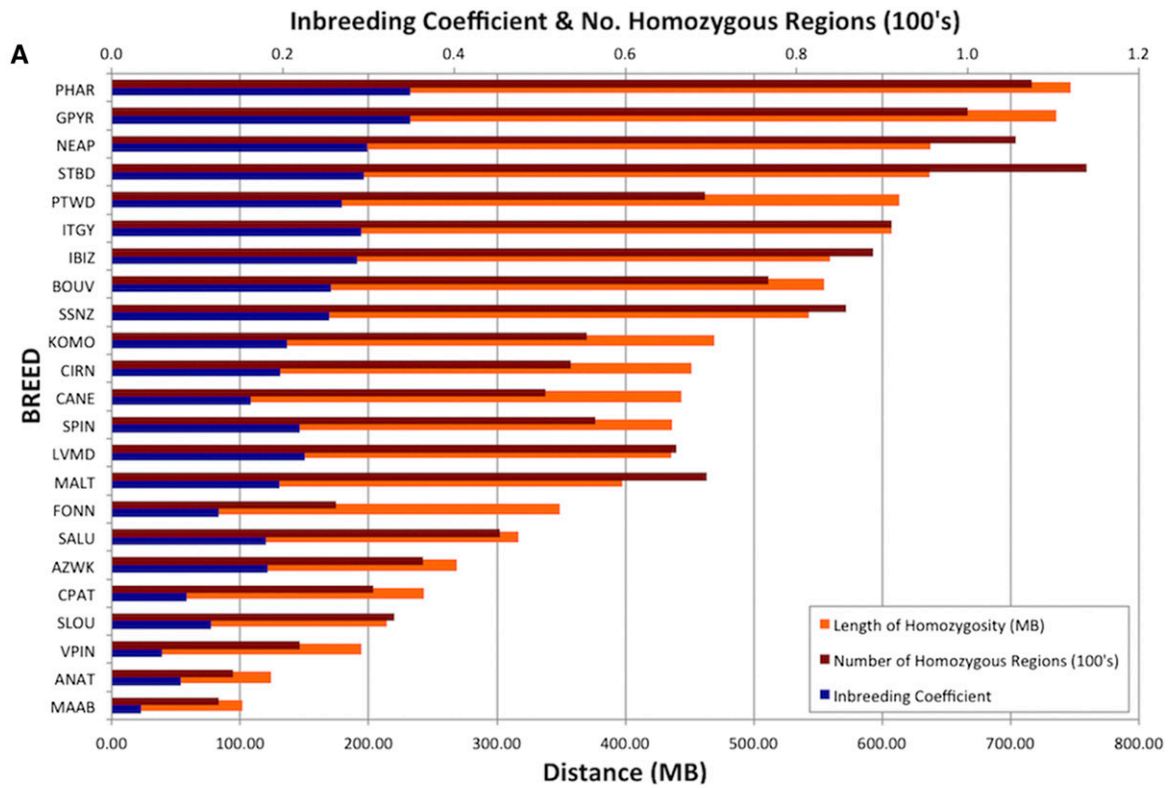
Estimated  $N_e$  was calculated for 15 breeds for which four or more dogs had SNP genotypes, over a timeframe of 13–995 previous generations (Figure 3). The most recent  $N_e$  values, at ~13 generations ago, ranged from 71 (Neapolitan Mastiff) to 303 (Saluki), with an across-breed mean of 117. The Fonní's Dog has a 13 generation  $N_e$  of 113. Each breed presented a rapidly decreasing  $N_e$  with a breed-specific mean  $\Delta N_e$  ranging from 81 (Saint Bernard) to 176 (Azawakh). The Fonní's Dog displayed a comparable mean  $\Delta N_e$  of 153.

### *Individual and expected breed range of individual variation in WGS*

We next sought to compare the level of unique variation across individuals representative of the Mediterranean dog breeds to assess the distinctiveness of each breed. A substantially deflated value of individual variation, relative to the levels in comparable breeds, would imply increased similarity to one or more of the other breeds. Individual variants were identified for each dog relative to a pool of 13 Mediterranean-breed dogs (Table 1). The resulting variants were classified as either the heterozygous or homozygous state by VCFtools. Variants found in the homozygous state in a single dog are considered representative of breed-specific variants with respect to the 13 Mediterranean breeds. However, variants identified in the heterozygous state in a single dog may represent breed-specific, though not breed-fixed, variants, or variants unique to the individual. The number of homozygous individual variants ranged from 6,776 (Spanish Water Dog) to 26,623 (Saluki) (Figure 2). The Saluki had the highest number of total individual variants (151,426) and the Istrian Shorthaired Hound had the lowest number (59,838). The across-breed mean for all individual variants was 82,056, with a mean of 15,203 for the homozygous variants (Table 3). The Fonní's Dog had 86,166 individual variants, with 9,224 present in the homozygous state.

### *Underlying population structure of Mediterranean breeds*

Population structure was calculated using the SNP data over a range of 2 through 20 possible groupings using all Mediterranean



**Figure 2** Quantitation of genetic measures of (A) homozygosity and inbreeding from SNP-chip data representing within-breed means, and (B) individual variation from WGS of individual dogs.

**Table 2 Inbreeding and homozygosity metrics from SNP-chip analyses, sorted by breed mean length of homozygosity**

Breed	No.	Inbreeding coefficient			Regions of homozygosity			Length of homozygosity (Mb)		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
MAAB	2	0.03	0.02	0.05	12.50	10	15	101.79	68.19	135.38
ANAT	6	0.08	0.05	0.15	14.17	7	26	124.00	387.46	301.20
VPIN	4	0.06	0.01	0.16	22.00	17	30	194.43	39.44	413.17
SLOU	5	0.12	0.02	0.19	33.00	6	61	213.91	16.29	371.94
CPAT	2	0.09	0.03	0.15	30.50	20	41	243.39	72.46	414.32
AZWK	5	0.18	0.13	0.36	36.40	26	52	268.91	91.20	644.37
SALU	19	0.18	0.02	0.37	45.32	4	83	316.90	9.24	729.28
FONN	6	0.13	0.01	0.30	26.17	8	49	348.71	59.55	760.59
MALT	10	0.20	0.10	0.34	69.50	38	96	397.53	208.98	748.85
LVMD	2	0.23	0.19	0.27	66.00	54	78	435.95	417.33	454.56
SPIN	2	0.22	0.20	0.24	56.50	51	62	436.80	381.50	492.11
CANE	9	0.16	0.02	0.42	50.67	16	75	443.81	32.00	1077.38
CIRN	5	0.20	0.08	0.38	53.60	34	85	451.72	133.27	906.52
KOMO	2	0.20	0.16	0.25	55.50	53	58	469.07	343.57	594.58
SSNZ	10	0.25	0.18	0.39	85.80	75	102	543.17	372.34	819.58
BOUV	8	0.26	0.14	0.47	76.75	57	89	554.60	257.60	1003.20
IBIZ	10	0.29	0.17	0.41	89.00	81	103	559.21	302.30	857.09
ITGY	10	0.29	0.20	0.42	91.10	69	104	607.67	411.60	884.78
PTWD	10	0.27	0.19	0.41	69.30	57	91	613.54	417.86	925.15
STBD	10	0.29	0.23	0.33	113.90	100	124	636.82	481.11	729.25
NEAP	6	0.30	0.26	0.36	105.67	98	119	637.51	507.46	818.94
GPYR	10	0.35	0.19	0.50	100.00	75	114	735.75	340.63	1096.51
PHAR	2	0.35	0.34	0.35	107.50	105	110	747.11	729.40	764.81

Min, minimum; max, maximum.

breeds, and 2 through 12 groupings for the Italian breeds. Maximum likelihood analyses identified  $K = 15$  (marginal likelihood =  $-1.00$ , variance =  $1.23 \times 10^{-4}$ ,  $\Delta K_{14-15} = -1.75 \times 10^{-3}$ ,  $\Delta K_{15-16} = 1.99 \times 10^{-3}$ ) and  $K = 3$  (marginal likelihood =  $-1.05$ , variance =  $1.86 \times 10^{-5}$ ,  $\Delta K_{2-3} = -3.82 \times 10^{-4}$ ,  $\Delta K_{3-4} = 4.85 \times 10^{-4}$ ) as the statistically appropriate number of structural groupings for Mediterranean and Italian breeds, respectively (Figure 4). This analysis grouped the Mediterranean breeds, as well as the Portuguese Water Dog, Bouvier des Flandres, Neapolitan Mastiff, Great Pyrenees, Ibizan Hound, Italian Greyhound, Maltese, Saint Bernard, and Standard Schnauzer as distinct breed clusters. The Saluki primarily formed a single grouping (orange in Figure 4), with some individuals demonstrating variable levels of a common multi-breed signature (dark purple), presumably related to population substructure resulting from the region of sample collection. The Pharaoh Hound shared 59% identity with the Ibizan Hound (green), and 8–12% identity with each of Neapolitan Mastiff (dark blue), Portuguese Water Dog (pale blue), Italian Greyhound (pale purple), and Saluki. The common signature (dark purple) accounted for the majority of the Volpino Italiano, Komondor, Spinone Italiano, Fonnì's Dog, Mastino Abruzzese, Cane Paratore, Levriero Meridionale, Azawakh, and Sloughi. The Neapolitan Mastiff structure represented 4–7% of the Cane Paratore; 2–5% of the Mastino Abruzzese; up to 7% of the Fonnì's Dog; and <2% of the Spinone Italiano, Komondor, and Volpino Italiano. The common multi-breed signature appeared in individual dogs of otherwise distinct breeds at a rate of 13% (Standard Schnauzer) to 30% (Maltese and Great Pyrenees). While the Anatolian Shepherd and Sloughi were comprised of the Saluki (orange) signature at a level of

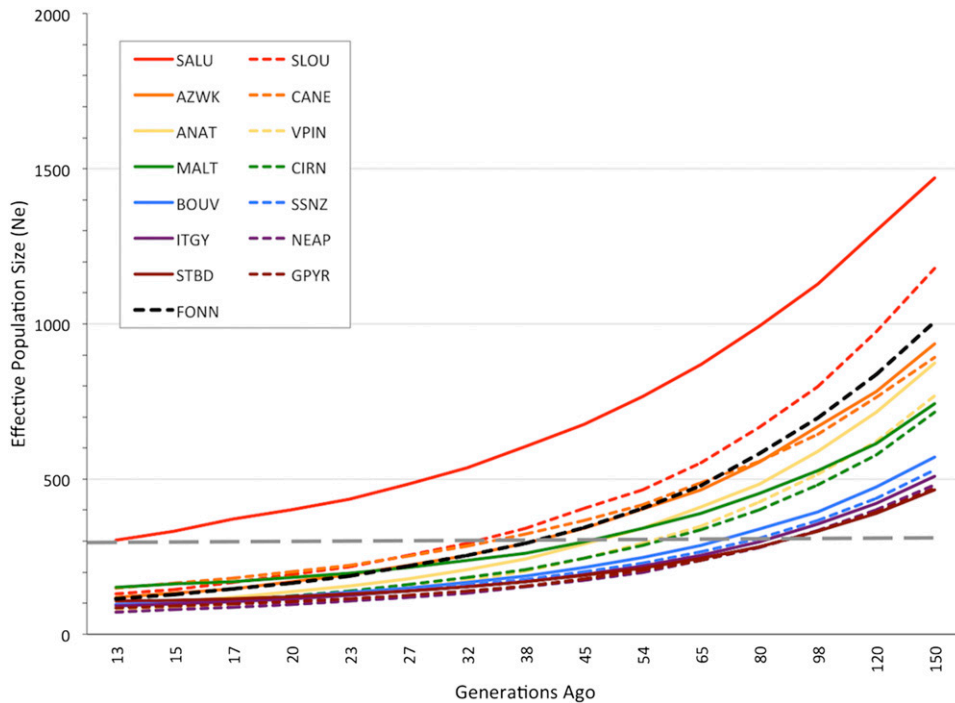
8–40%, the remaining identity was that of the common multi-breed cluster.

Within the Italian breeds ( $K = 3$ ) the Maltese and Neapolitan Mastiff identify as single clusters, with the Cirneco dell'Etna, Cane Paratore, Fonnì's Dog, Komondor, Levriero Meridionale, Mastino Abruzzese, Volpino Italiano, Pharaoh Hound, and Spinone Italiano combined to produce the third cluster. The Cane Corso was comprised of 27–73% of the Neapolitan Mastiff signature, with the remainder being the multi-breed cluster. A total of 2 of the 10 Maltese demonstrated 7 or 27% identity with the multi-breed grouping.

Interbreed genome-wide  $F_{ST}$  values range from 0.0167 (Fonnì's Dog  $\times$  Mastino Abruzzese) to a maximum of 0.2942 (Pharaoh Hound  $\times$  Neapolitan Mastiff) (Figure 5). Values closer to zero highlight populations that are comparatively more genetically homogeneous, while values closer to one reflect populations with greater genetic divergence. In this analysis, the Fonnì's Dog is less diverged from the Mastino Abruzzese ( $F_{ST} = 0.0167$ ) and the Cane Paratore ( $F_{ST} = 0.0284$ ), compared to the Saint Bernard ( $F_{ST} = 0.1669$ ) and the Neapolitan Mastiff ( $F_{ST} = 0.1591$ ). The distance between the Fonnì's Dog and the Neapolitan Mastiff is unexpected due to the Italian heritage of both breeds.

#### **Genetic breed variation visualized through PCA and phylogeny**

PCA was conducted using two separate breed groupings: the Mediterranean and Italian groups. Analysis of the first four principal components (PCs) (PC1 = 5.54,  $P = 1.33 \times 10^{-46}$ ; PC2 = 4.89,  $P = 5.68 \times 10^{-42}$ ; PC3 = 4.70,  $P = 5.68 \times$



**Figure 3**  $N_e$  estimates from SNP-based calculation of LD.

$10^{-51}$ ;  $PC4 = 4.49$ ,  $P = 2.78 \times 10^{-62}$ ) of the Mediterranean breeds (Figure 6, A and B) showed that the Saint Bernard, Italian Greyhound, and Great Pyrenees cluster as distinct breeds, independent of the others. The Neapolitan Mastiff and Cane Corso identify as separate, tightly clustered groupings, as did the Standard Schnauzer and Portuguese Water Dog. The remaining Mediterranean breeds appeared to form loose breed clusters that were not readily distinguishable from one another. The Italian group subset improved resolution between those breeds ( $PC1 = 3.84$ ,  $P = 2.41 \times 10^{-27}$ ;  $PC2 = 2.71$ ,  $P = 1.40 \times 10^{-23}$ ;  $PC3 = 2.03$ ,  $P = 1.83 \times 10^{-13}$ ;  $PC4 = 1.65$ ,  $P = 1.88 \times 10^{-05}$ ) (Figure 6, C and D). The Cane Corso and Neapolitan Mastiff continued to form distinct clusters in close proximity to each other. The Maltese formed a clear cluster, and the Pharaoh Hound and Cirneco dell'Etna grouped together. However, the Cane Paratore, Fonní's Dog, Komondor, Levriero Meridionale, Mastino Abruzzese, Spinone Italiano, and Volpino Italiano did not resolve into distinct breed clusters. Further reduction of these regional breeds (Figure 6, E and F) resolved  $PC1-3$  into breed-specific clusters, though without statistically significant separation ( $PC1 = 1.51$ ,  $P = 0.510$ ;  $PC2 = 1.46$ ,  $P = 0.319$ ;  $PC3 = 1.40$ ,  $P = 0.185$ ). The Spinone Italiano and Levriero Meridionale were most divergent from the others. The Mastino Abruzzese, Cane Paratore, and Komondor breeds were distinct, yet closely related, groups. The Volpino Italiano and Fonní's Dog, while also distinct groupings, were more diffuse.

Similar phylogenetic relationships were also observed among the same dogs (Figure 7). As with the PCA, the cladograms indicated that the Cane Corso and Neapolitan Mastiff, and the Standard Schnauzer and Portuguese Water Dog, are closely related. However, the common structural group-

ing, visualized using analysis with the STRUCTURE program (purple in Figure 4), encompassing the Cane Paratore, Mastino Abruzzese, Fonní's Dog, Spinone Italiano, Komondor, Volpino Italiano, Saluki, Anatolian Shepherd, Sloughi, Azawakh, Levriero Meridionale, and Cirneco dell'Etna, separated in the cladograms in a manner that best reflected the morphological or ancestral commonalities between breeds. For instance, the Cirneco dell'Etna, Pharaoh Hound, and Ibizan Hound, and the Sloughi, Levriero Meridionale, and Azawakh formed two monophyletic groups in putative sighthound clades. The Fonní's Dog, Mastino Abruzzese, Cane Paratore, and Volpino Italiano, were monophyletic and branched immediately outside of the Portuguese Water Dog and Standard Schnauzer; breeds that reflect greater phenotypic similarity with each other than with either the sighthounds or the mastiffs (Neapolitan Mastiff, Cane Corso, and Saint Bernard).

#### **Admixture between Fonní's Dogs and Mediterranean breeds**

We next focused on the Fonní's Dog in terms of its historical involvement in breed development throughout the Mediterranean. Breed admixture with the Fonní's Dog was calculated using three separate algorithms (Figure 8). Introgression of the Fonní's Dog, Portuguese Water Dog, and Cane Paratore was identified within the SNP data by the TreeMix three\_pop and AdmixTools software programs. TreeMix predicted a 17.43% contribution of Fonní's Dog to the ancestor of the Portuguese Water Dog and Cane Paratore ( $P = 0.00723$ ). Patterson's  $D$  from AdmixTools replicated these findings for Fonní's Dog and Portuguese Water Dog with  $D(O, Portuguese Water Dog; Fonní's Dog, X) = -0.0139$  to  $-0.1287$ ,  $Z = -3.737$  to  $-25.274$ ,  $D(O, Fonní's Dog; Portuguese Water$



**Table 3 WGS individual variants for Mediterranean breed dogs**

Dog	Heterozygotes	Homozygotes	Total
Istrian Shorthaired Hound	49,078	10,760	59,838
Lagotto Romagnolo	53,820	11,733	65,553
Berger Picard	44,080	25,049	69,129
Italian Greyhound	53,206	16,827	70,033
Great Pyrenees	54,289	22,677	76,966
Standard Schnauzer	61,347	16,572	77,919
Spanish Galgo	61,434	17,335	78,769
Saint Bernard	65,669	13,884	79,553
Spanish Water Dog	75,597	6,776	82,373
Cane Corso	72,221	12,243	84,464
Portuguese Water Dog	76,602	7,939	84,541
Fonni's Dog	76,942	9,224	86,166
Saluki	124,803	26,623	151,426

Dog,  $X$ ) =  $-0.0127$  to  $-0.1309$ ,  $Z = -3.078$  to  $-25.797$ , and  $D(O, \text{Fonni's Dog}; X, \text{Portuguese Water Dog}) = 0.0100$  to  $0.0767$ ,  $Z = 3.016$  to  $22.189$ . Likewise, AdmixTools reiterated the introgression of Fonni's Dog and Cane Paratore with  $D(O, \text{Cane Paratore}; \text{Fonni's Dog}, X) = -0.0104$  to  $-0.1317$ ,  $Z = -3.108$  to  $-26.025$ ,  $D(O, \text{Fonni's Dog}; \text{Cane Paratore}, X) = -0.0096$  to  $-0.1289$ ,  $Z = -3.055$  to  $-25.805$ , and  $D(O, \text{Fonni's Dog}; X, \text{Cane Paratore}) = 0.0104$ – $0.0811$ ,  $Z = 3.643$ – $22.974$ .

A second instance of introgression was identified between Fonni's Dog and Komondor within the WGS dataset, with ANGSD-calculated significant  $D$ -statistics for  $D(O, \text{Fonni's Dog}; \text{Komondor}, \text{Cane Corso}) = 0.039$ ,  $Z = 4.366$ ; and  $D(O, \text{Komondor}; \text{Fonni's Dog}, \text{Cane Corso}) = 0.031$ ,  $Z = 3.920$ ; and with nonsignificant values of  $D(O, \text{Cane Corso}; \text{Komondor}, \text{Fonni's Dog}) = 0.008$ ,  $Z = 1.028$ . This was confirmed with AdmixTools, which revealed a significant  $D$  from SNP data:  $D(O, \text{Fonni's Dog}; \text{Komondor}, X) = -0.0116$  to  $-0.1123$ ; and  $D(O, \text{Fonni's Dog}; X, \text{Komondor}) = 0.0116$ – $0.1280$ ,  $Z = 3.142$ – $25.720$ .

Introgression between Fonni's Dog and Saluki was also observed using ANGSD:  $D(O, \text{Fonni's Dog}; \text{Saluki}, \text{Cane Corso}) = 0.033$ ,  $Z = 3.893$ ;  $D(O, \text{Saluki}; \text{Fonni's Dog}, \text{Cane Corso}) = 0.040$ ,  $Z = 5.075$ ; and  $D(O, \text{Cane Corso}; \text{Saluki}, \text{Fonni's Dog}) = -0.007$ ,  $Z = -1.060$ . These results were replicated with SNP data using AdmixTools, resulting in  $D$ -statistics of  $D(\text{Saluki}, X; \text{Fonni's Dog}, Y) = 0.007$ – $0.131$ ,  $Z = 3.047$ – $29.352$ .

Separately, TreeMix predicted phylogeny trees for the Mediterranean breeds with allowance for 1 through 10, 15, and 25 introgression events. The  $\ln(\text{likelihood})$ , indicating the probability of the suggested relationships explaining the genomic data, for the predicted phylogenies increased with additional allowed introgressions, from  $\ln(\text{likelihood})_1 = 1599.06$  and  $\ln(\text{likelihood})_{25} = 1888.17$ . With the inclusion of 25 introgression events there was not yet any indication of decreasing  $\ln(\text{likelihood})$  and none of the proposed introgressions involved the Fonni's Dog (Figure 8).

### Regions of homozygosity represent putative regions of selection

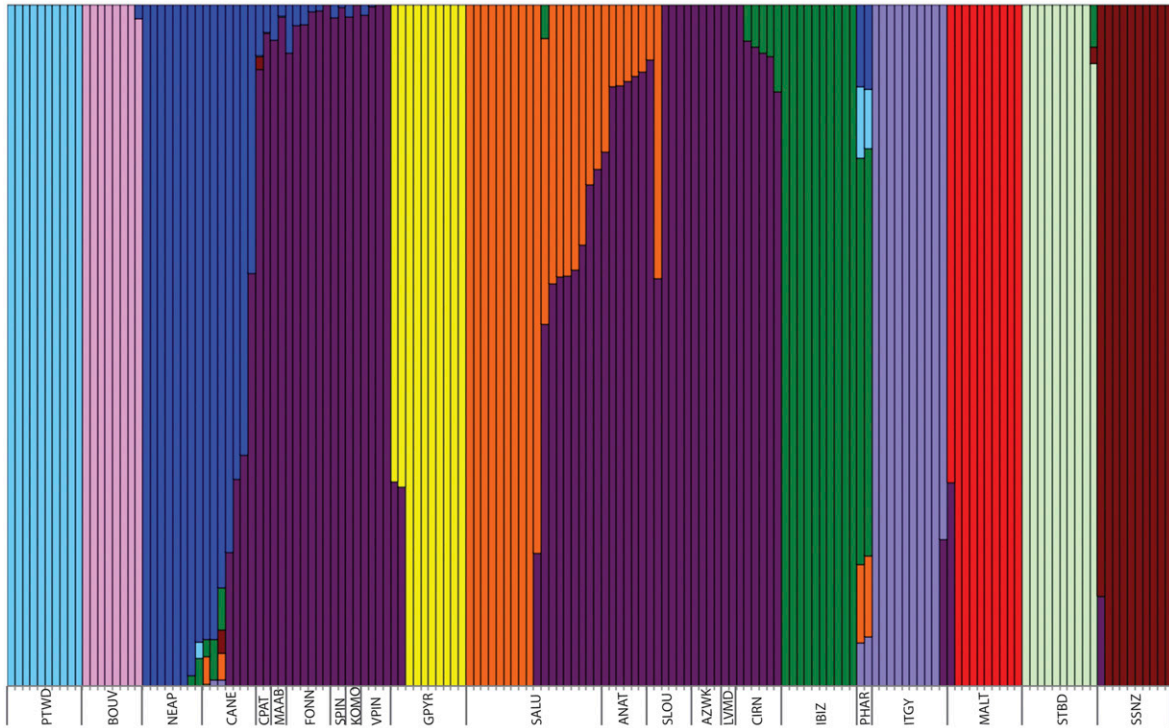
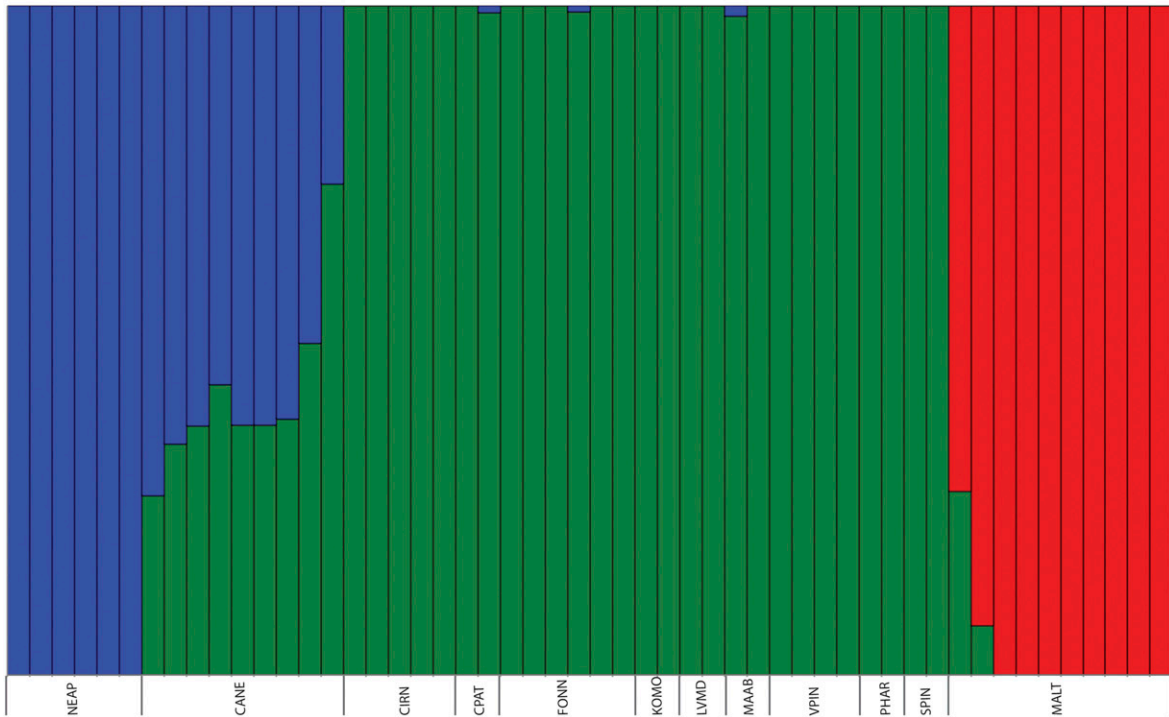
To identify regions of the genome potentially under selection in the Fonni's Dog, we calculated SNP genotypes for each of

the six Fonni's Dogs. This revealed 258 total regions of homozygosity (range per dog = 25–61), each of which spanned a minimum of 10 sequential SNP-chip variants. Seven regions were shared across four of the six dogs. Analysis of SNP haplotypes across each shared region revealed 11 even shorter regions, defined by three or more sequential SNPs that were shared across all six Fonni's Dogs (Table 4). Individual dog variants extracted from the WGS sequence were filtered for the 11 SNP-based regions of shared homozygosity, resulting in 9 heterozygous and 1 homozygous variant within the selected regions. Each of these variants was located in a noncoding region and annotated as a modifier by SNPeff. (Cingolani *et al.* 2012).

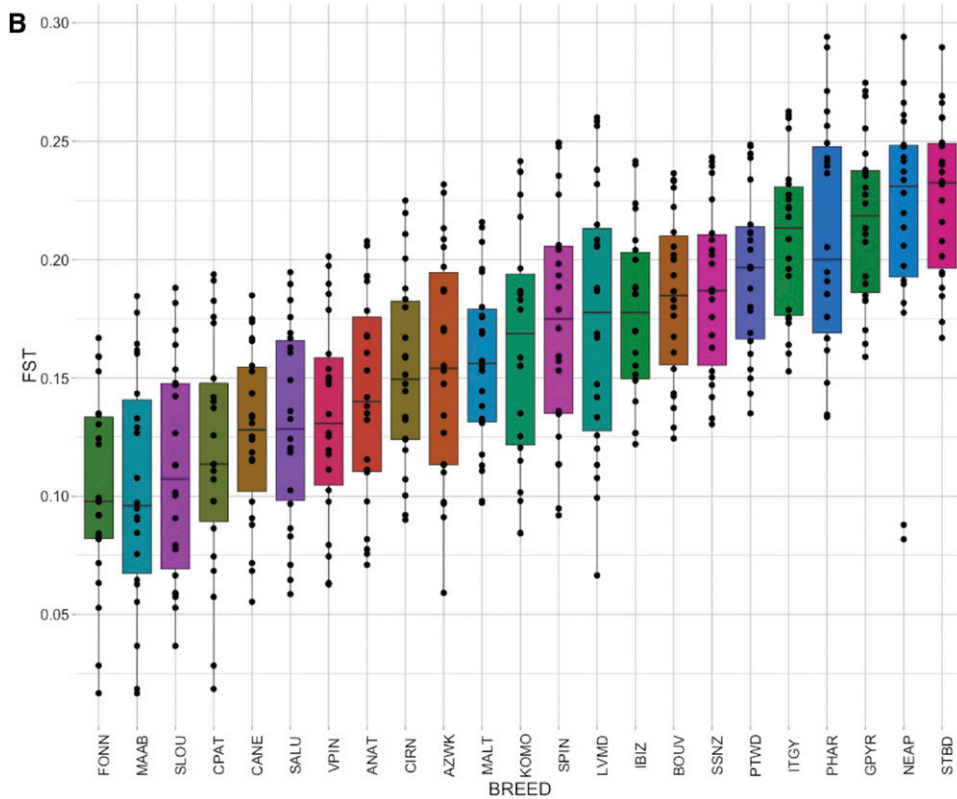
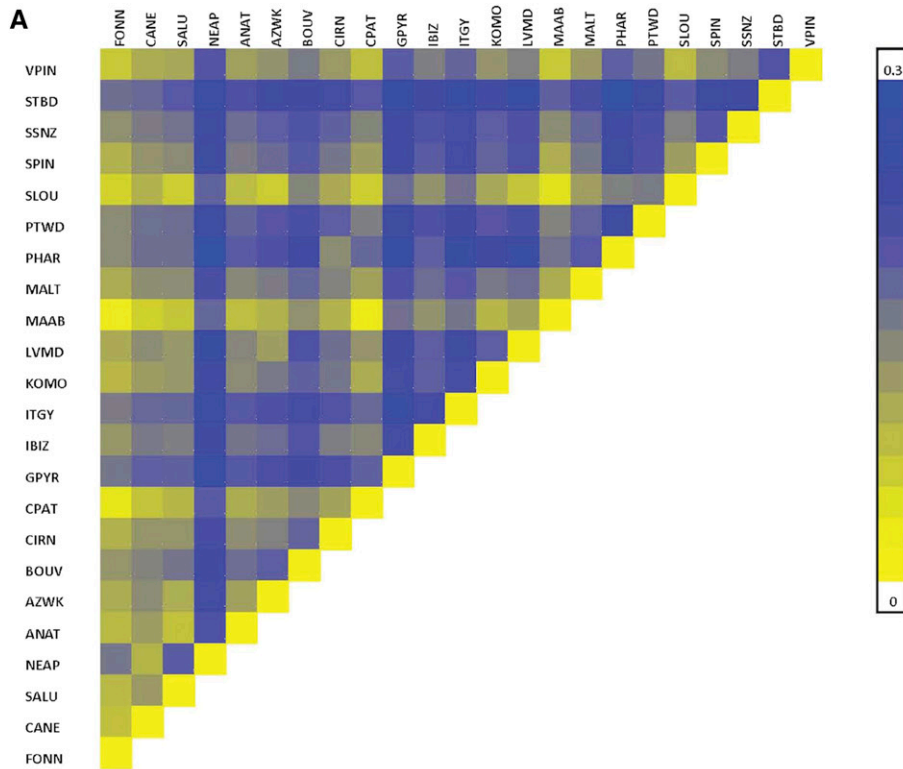
### Discussion

Genetic investigation of population isolates can provide insights into inheritance of both rare and complex traits. In humans, for instance, studies of Bedouin tribes have successfully localized causal loci for single gene traits such as Bardet-Biedl Syndrome 3 (Frag and Teebi 1989; Sheffield *et al.* 1994; Chiang *et al.* 2004), nonsyndromic hearing loss (Scott *et al.* 1996), and infantile nephronophthisis (Haider *et al.* 1998). Studies of the Finnish population, by comparison, have been more useful for studies of complex traits, revealing genes that increase susceptibility to various cancers (Nyström-Lahti *et al.* 1994; Kainu *et al.* 2000; Sarantaus *et al.* 2000; Baffoe-Bonnie *et al.* 2005; Hartikainen *et al.* 2005; Rokman *et al.* 2005). Icelandic populations have also been used successfully to map risk alleles associated with a variety of complex traits and psychiatric conditions (Hicks *et al.* 2002; Thorgeirsson *et al.* 2003; Karason *et al.* 2005; Arason *et al.* 2010).

Studies of geographically or culturally restricted breeding in human populations have also informed our understanding of movement and growth of human populations and, with them, genes of interest. For instance, characterization of point mutations associated with  $\beta$ -thalassemia in Kurdistan Jewish populations have identified regional admixture and founder effects specific to isolated Kurdish subpopulations (Rund *et al.* 1991). Geographically remote populations can likewise reveal historical human migration or settlement events, exemplified by the Pitcairn Island population which we now know has maternal Tahitian and paternal European lineages (Benton *et al.* 2015). Within-country cultures that restrict marriage outside the community have proven similarly informative (Ginns *et al.* 1998; Stone *et al.* 1998; Hsueh *et al.* 2003; Seboun *et al.* 2005; Simpson *et al.* 2009; Hou *et al.* 2013; Georgi *et al.* 2014; Kember *et al.* 2015). The success of using population isolates for mapping of complex and rare disorders lies predominantly in the common features of such isolates: decreased genomic diversity, (Sajantila *et al.* 1996; Capocasa *et al.* 2014; Di Gaetano *et al.* 2014), increased LD (Bendjilali *et al.* 2014), and increased inbreeding values (Zhai *et al.* 2016). Dog breeds reflect these same characteristics, resulting in each breed forming a distinct population

**A****Mediterranean Breeds, K = 15****B****Italian Breeds, K = 3**

**Figure 4** STRUCTURE analysis of (A) Mediterranean and (B) Italian dog breeds. Colors indicate separate STRUCTURE group signatures for individual dogs. Maximum likelihood predicts groupings of  $K = 15$  and  $K = 3$  for Mediterranean and Italian breeds, respectively.

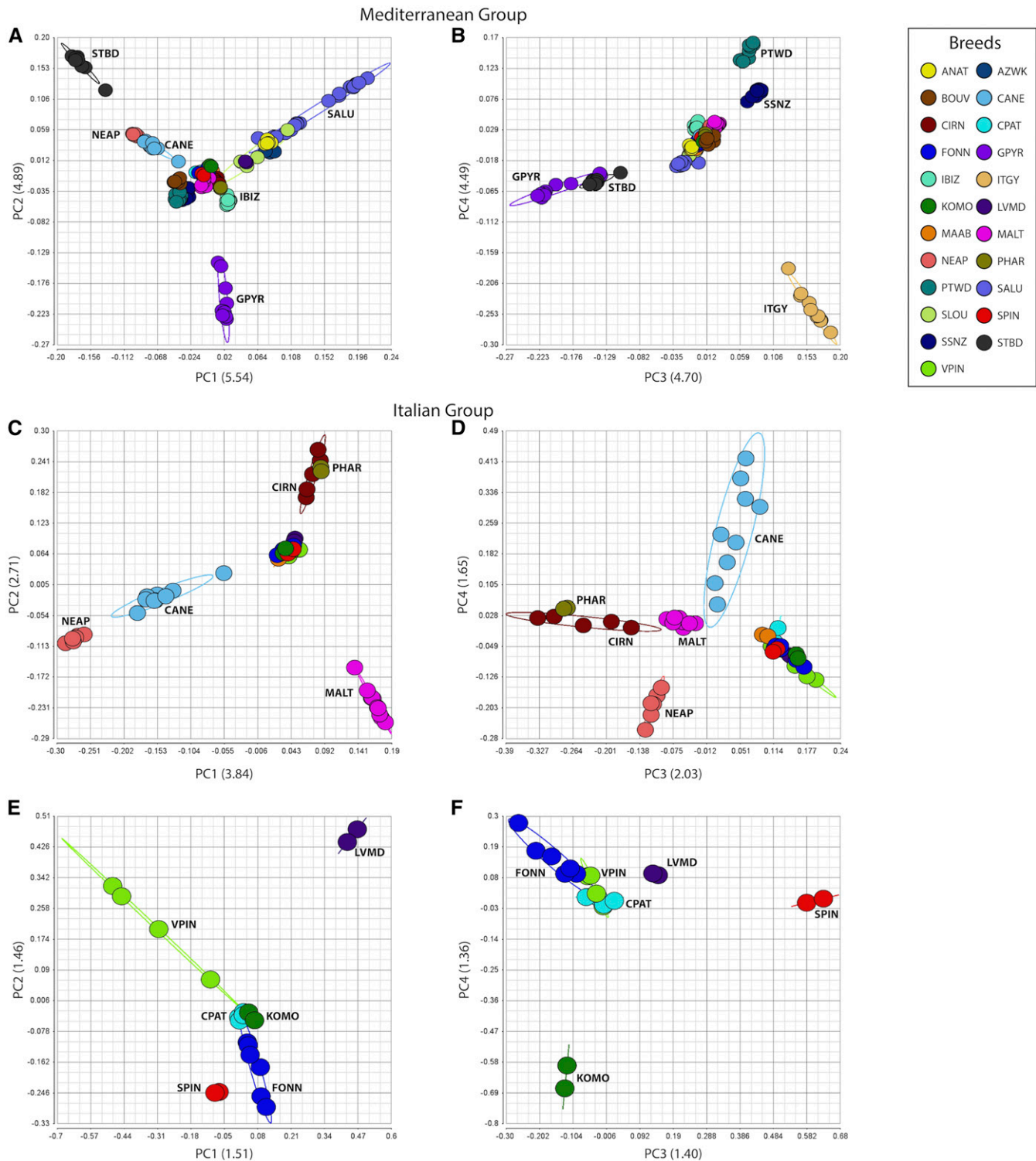


**Figure 5** Weir and Cockerham's weighted  $F_{ST}$  values based on breed-to-breed comparisons of SNP genotypes. (A) Scores of 0 (yellow) indicate perfect identity, scores of 0.3 (blue) indicate the highest level of divergence observed in this set of breeds. (B) Distribution of breed-specific  $F_{ST}$  values.

isolate formed through human intervention and selection (reviewed in Boyko 2011; Schoenebeck and Ostrander 2014).

Previous molecular analyses of the Fonn's Dog of Sardinia provide strong evidence that this niche population meets the

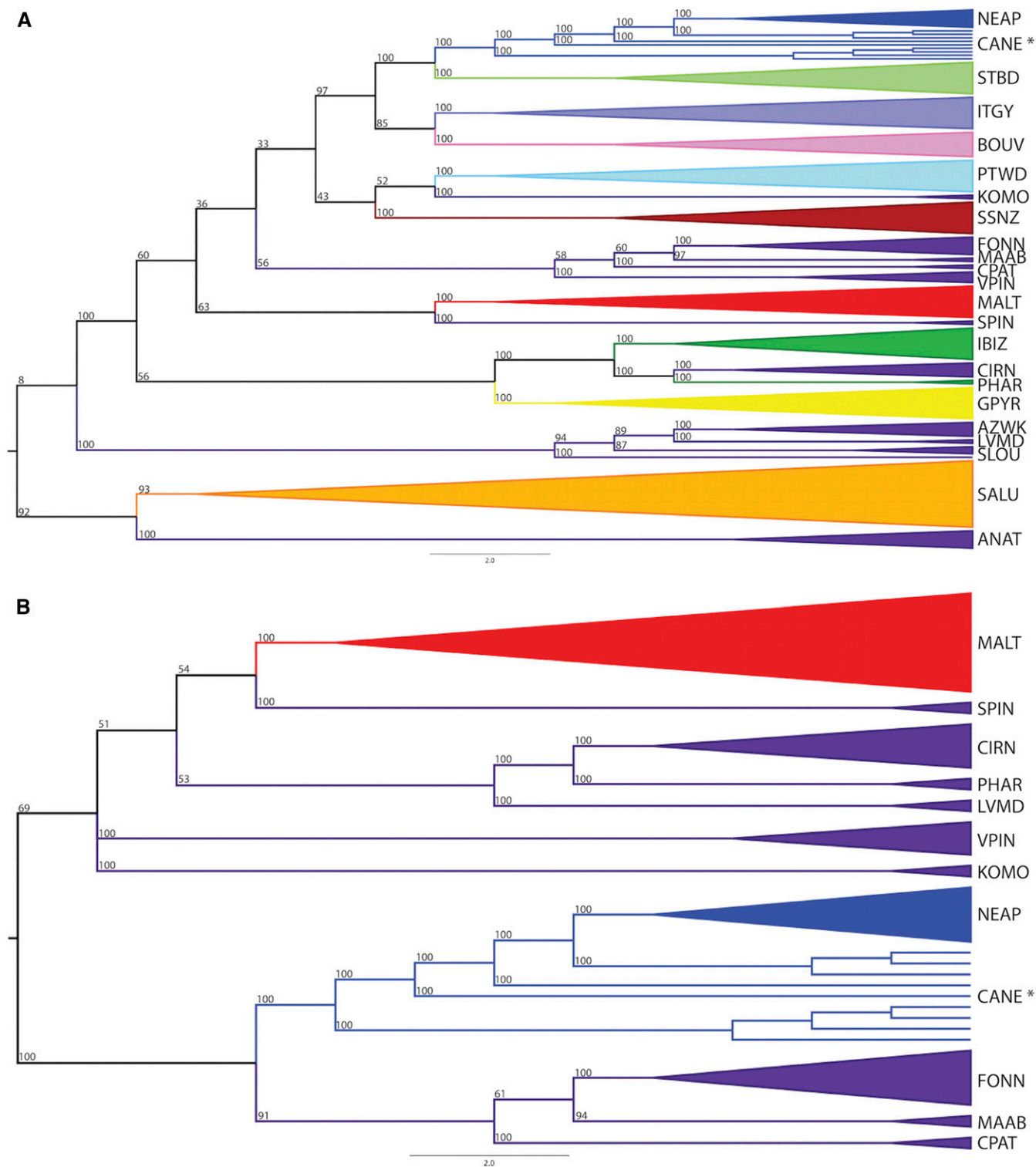
formal definition of a breed (Sechi *et al.* 2016). It appears genetically distinct from other breeds in the region as measured by a small set of microsatellite-based polymorphic markers, and both physical and behavioral traits are retained



**Figure 6** PCA of SNP-chip data for (A and B) Mediterranean breeds (PC1 = 5.54,  $P = 1.33 \times 10^{-46}$ ; PC2 = 4.89,  $P = 5.68 \times 10^{-42}$ ; PC3 = 4.70,  $P = 5.68 \times 10^{-51}$ ; PC4 = 4.49,  $P = 2.78 \times 10^{-62}$ ), (C and D) all Italian breeds (PC1 = 3.84,  $P = 2.41 \times 10^{-27}$ ; PC2 = 2.71,  $P = 1.40 \times 10^{-23}$ ; PC3 = 2.03,  $P = 1.83 \times 10^{-13}$ ; PC4 = 1.65,  $P = 1.88 \times 10^{-05}$ ), and (E and F) regional Italian breeds (PC1 = 1.51,  $P = 0.51$ ; PC2 = 1.46,  $P = 0.32$ ; PC3 = 1.40,  $P = 0.19$ ; PC4 = 1.40,  $P = 0.04$ ). Breed clusters are denoted with ellipses of 2 SD. Breed abbreviations are listed in Table 1.

from generation to generation (Sechi *et al.* 2016). Historically, microsatellite-based studies have proven effective for differentiating between dog breeds (Parker *et al.* 2004; Leroy

*et al.* 2009; Mellanby *et al.* 2013), identifying varieties that inhabit a common insular region (Parra *et al.* 2008; Pribanova *et al.* 2009; Suarez *et al.* 2013), or identifying a

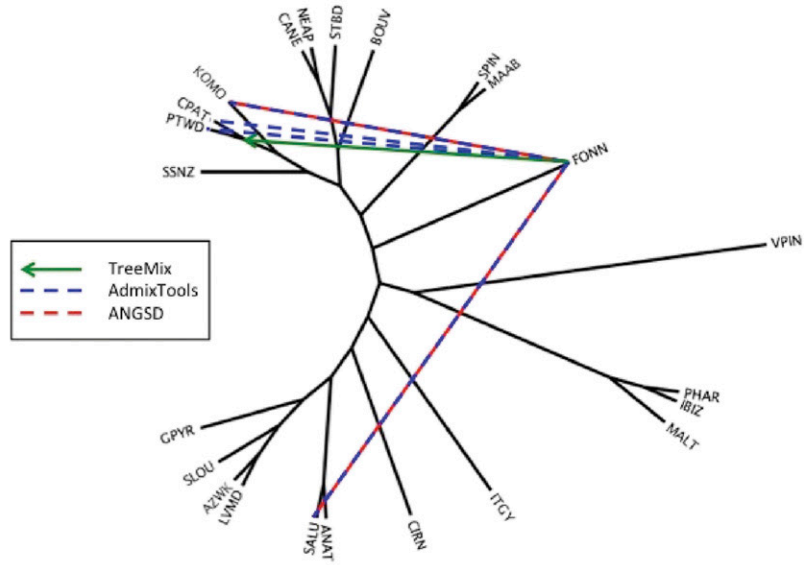


**Figure 7** SNP-based neighbor-joining cladograms for (A) Mediterranean and (B) Italian breeds. Clade coloration reflects the designations assigned by FastSTRUCTURE (Figure 4). Branch numbers indicate the bootstrapping values from 100 repetitions. \*, the nine Cane Corso dogs did not form a single clade, however the bootstrap values for their phylogenetic locations relative to each other are each 100. Breed abbreviations are listed in Table 1.

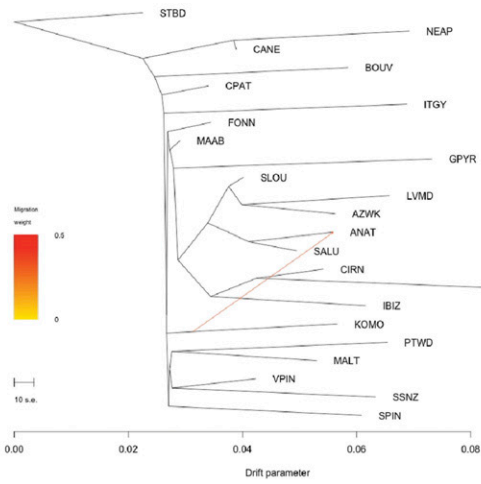
subset of the component breeds that contributed to individual mixed breed dogs (Parker *et al.* 2004). However, since microsatellites are multi-allelic and highly mutable, they are useful largely for family studies and not necessarily popula-

tion studies, providing information that reflects decades and not hundreds or thousands of years. When indels and copy number variants are also considered, it is clear that canine genetic studies are best served using dense SNP-chip analyses

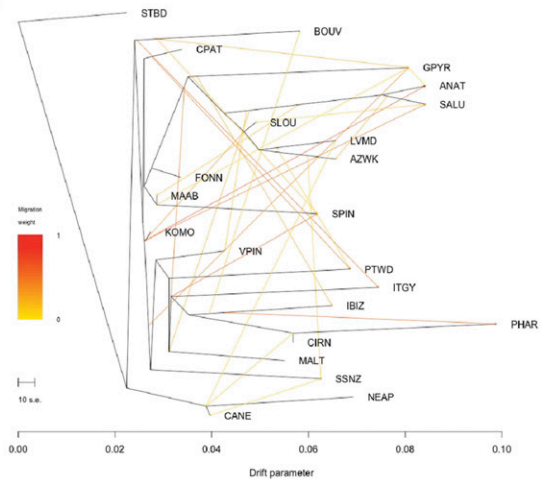
**A**



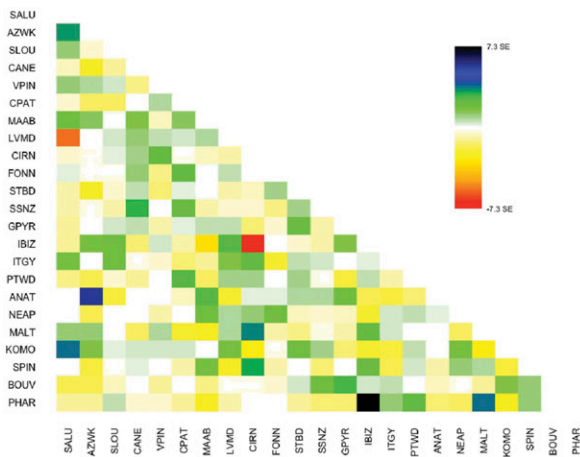
**B**



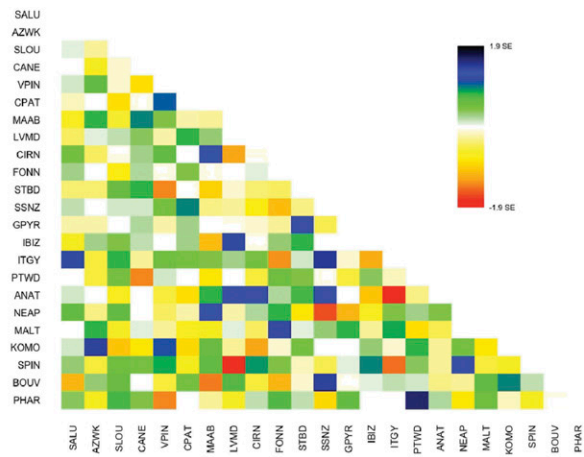
**C**



**D**



**E**



**Figure 8** (A) Admixture predictions by three computational algorithms. TreeMix three\_pop analysis (green arrow) identified Fonní's Dog contribution to the Portuguese Water Dog/Cane Paratore clade at 17.43% with  $P = 0.00723$ . ANGSD (red dashed line) calculated significant values of  $D(O, \text{Fonní's Dog; Komondor, Cane Corso}) = 0.039$ ,  $Z = 4.366$ ;  $D(O, \text{Komondor; Fonní's Dog, Cane Corso}) = 0.031$ ,  $Z = 3.920$ ;  $D(O, \text{Fonní's Dog; Saluki, Cane Corso}) = 0.033$ ,  $Z = 3.893$ ;  $D(O, \text{Saluki; Fonní's Dog, Cane Corso}) = 0.040$ ,  $Z = 5.075$ . AdmixTools (blue dashed line) supports the findings of TreeMix with significant  $D$ -statistic values for Fonní's Dog and Portuguese Water Dog of  $D(O, \text{Portuguese Water Dog; Fonní's Dog, } X) = -0.0139$  to  $-0.1287$ ,

**Table 4** Regions of homozygosity shared by six Fonní's Dogs, as identified from SNP-chip analysis

Region of homozygosity	Length (bp)	Personal variants <sup>a</sup>	Variant impact <sup>b</sup>	Gene affected
chr6:2,942,512-2,998,885	56,373	None		
chr7:551,194-675,969	124,775	None		
chr7:733,453-914,233	180,780	chr7:742,610	Modifier	<i>LGR6</i>
chr7:928,930-988,934	60,004	chr7:955,472	Modifier	<i>ELF3</i>
chr7:1,027,580-1,298,049	207,469	chr7:1,179,044	Modifier	<i>NAV1</i>
		chr7:1,190,728	Modifier	<i>NAV1</i>
		chr7:1,221,667	Modifier	<i>NAV1</i>
		None		
chr8:1,069,127-1,108,694	39,567	None		
chr13:991,157-1,022,654	31,497	None		
chr13:1,115,146-1,156,529	41,383	chr13:1,131,868	Modifier	<i>VPS13B</i>
		chr13:1,157,079	Modifier	<i>VPS13B</i>
chr13:1,272,226-1,291,574	19,348	chr13:1,290,464	Modifier	<i>VPS13B</i>
chr13:1,571,048-1,655,073	84,025	chr13:1,572,305	Modifier	<i>VPS13B</i>
chr22:2,578,384-2,618,437	40,053	chr13:2,584,594	Modifier	<i>FNDC3A</i>

Chr, chromosome.

<sup>a</sup> Individual variants from Fonní's Dog WGS that are within the homozygous regions.

<sup>b</sup> Variant impact and affected genes annotated from CanFam3.1.76.

combined with WGS. Such methods offer the ability to use bioinformatic tools developed for problems such as population structure, a common issue in dog breeds and one for which it is nearly impossible to correct for using microsatellites, thus providing a degree of confidence and precision not otherwise available (Ostrander *et al.* 1993; Francisco *et al.* 1996).

The Fonní's Dog represents an unusual case study in that it maintains a standard appearance and well-defined behavior in the absence of an organized, human-directed breeding program. Such programs are the hallmark of most modern breeds and have been key to both the development and maintenance of purebred dogs (American Kennel Club 2006). We chose, therefore, to further investigate the Fonní's Dog as a means to understand how an isolated population can be developed through selection based solely on functional aptitude, without dilution from surrounding mating populations, to meet environmental demands (Figure 1). We proposed to do this by assessing the historical development of the breed and the composition of modern canine genome structure using SNP and WGS analyses.

Accounts of the Fonní's Dog from the late 1880s describe them as "... a very large and shaggy haired race, half-mastiff half-bloodhound ..." (Tyndale 1849), and a combination of "... the greyhound with a big dog ... where the opposite dimensions of the greyhound and the mastiff cancel each other out" (Cetti 1774). While there is the expected level of husbandry by breed enthusiasts, selection of dogs for breeding purposes is based predominantly on functional ability, and there has only been a very limited organized attempt

to direct the breed or to select on the basis of appearance. The modern standard description of the breed emphasizes the fundamental hostility of Fonní's Dogs toward strangers, natural tendencies toward territoriality, and loyalty only to its charges, much as did early descriptions of the breed (Cetti 1774) (<http://www.canedifonni.it/web/en/the-official-standard/>). Thus, within the context of dog breeds existing in Italy 130 years ago, the Fonní's Dog had a noteworthy, although limited, appeal that it retains today.

To accomplish our goals, we used data from our laboratory that tested correlations between SNP-derived homozygosity values; SNP-, WGS-, and pedigree-derived inbreeding coefficients; and effective numbers of early breed ancestors inclusive of several breeds (unpublished data). This data revealed that increases in either the number or length of homozygous regions within a breed, as measured from SNP-chip data, significantly correlate with an increase in inbreeding coefficients calculated from SNP and WGS data, and also correlate with a decrease in the effective number of breed ancestors as observed by pedigree analysis (unpublished data). Applying those principles to the present study, we found that the sum length of homozygous stretches in the genome of Fonní's Dogs, the number of regions of homozygosity, and the inbreeding values for the Fonní's Dog were intermediate to those defining the remaining Mediterranean breeds (Figure 2). The Mediterranean breeds with lower homozygosity values than the Fonní's Dog, suggestive of a more complex ancestral history, include the Saluki, Azawakh, Cane Paratore, Sloughi, Volpino Italiano, Anatolian Shepherd, and Mastino Abruzzese. The relative ranking of the Fonní's Dog in

$Z = -3.737$  to  $-25.274$ ;  $D(O, \text{Fonní's Dog}; \text{Portuguese Water Dog}, X) = -0.0127$  to  $-0.1309$ ,  $Z = -3.078$  to  $-25.797$ ;  $D(O, \text{Fonní's Dog}; X, \text{Portuguese Water Dog}) = 0.0100$ – $0.0767$ ,  $Z = 3.016$ – $22.189$ ; Cane Paratore of  $D(O, \text{Cane Paratore}; \text{Fonní's Dog}, X) = -0.0104$  to  $-0.1317$ ,  $Z = -3.108$  to  $-26.025$ ;  $D(O, \text{Fonní's Dog}; \text{Cane Paratore}, X) = -0.0096$  to  $-0.1289$ ,  $Z = -3.055$  to  $-25.805$ ;  $D(O, \text{Fonní's Dog}; X, \text{Cane Paratore}) = 0.0104$ – $0.0811$ ,  $Z = 3.643$ – $22.974$ ; Komondor of  $D(O, \text{Fonní's Dog}; \text{Komondor}, X) = -0.0116$  to  $-0.1123$ ;  $D(O, \text{Fonní's Dog}; X, \text{Komondor}) = 0.0116$ – $0.1280$ ,  $Z = 3.142$ – $25.720$ ; and Saluki of  $D(\text{Saluki}, X; \text{Fonní's Dog}, Y) = 0.007$ – $0.131$ ,  $Z = 3.047$ – $29.352$ . TreeMix predicted phylogenies with (B) 1 or (C) 25 allowed introgression events, and corresponding standard error residuals for (D) 1 and (E) 25 introgressions.

terms of homozygosity and inbreeding coefficients confirms preliminary data (Sechi *et al.* 2016), categorizing Fomni's Dog as a distinctive dog breed compared to a dataset of dogs from the Mediterranean region. Additionally, the rate of  $\Delta N_e$  demonstrates that the Fomni's Dog obtained its breed status at a rate of population consolidation equivalent to breeds regulated by intensive selection systems.

We produced WGS from a single Fomni's Dog, sampled from the Cagliari region of Sardinia, achieving a mean depth of  $36\times$ . Assessment of WGS-derived individual variants in that dog, relative to a pool of 12 additional Mediterranean dogs, was undertaken to obtain a measure of individual variation for each dog compared to representatives of distinct breeds from the same geographic region. Specific to the Fomni's Dog were 86,166 variants, of which 9,224 were in the heterozygous state and 76,942 were in the homozygous state (Table 3). The Saluki displayed the greatest number of individual variants (151,426), with 65,260 more variants than the dog with the second-highest number of individual variants, the Fomni's Dog. However, excluding the Saluki, the 12 remaining dogs with WGS presented numbers of individual variation within a range of 26,328. Therefore, the level of individual variation in the Saluki is substantially greater than any of the other breeds. Populations that have experienced intensive natural or artificial selective pressures would be expected to have a lower number of individual variants in the heterozygous state than dogs from less well-defined breeds. Conversely, individual variants are more indicative of breed-specific, though not necessarily breed-fixed, privatization (Szpiech *et al.* 2013; Marsden *et al.* 2016).

Combined consideration of breed-specific measures of homozygosity, inbreeding,  $N_e$ , and individual variation further confirms our designation of dog breeds as population isolates. While the precise degrees of population homogeneity vary between other breeds, the Fomni's Dog has not demonstrated any substantial deviation from this definition of breed relative to the other Mediterranean breeds. We next addressed the observation that, contrary to most modern breeds, the Fomni's Dog has attained this breed status without the highly structured framework of selection toward a defined ideal.

To determine the foundational components of the Fomni's Dog, several analyses were undertaken. Genotype data from  $\sim 150,000$  SNPs were processed through the program STRUCTURE, as well as used in PCA, neighbor-joining phylogeny, and  $F_{ST}$  calculations. Each associated algorithm measures distinct population characteristics: model-based clustering is calculated from the program STRUCTURE, population differentiation from  $F_{ST}$ , definition of variable correlation from PCA, and visualization of genetic distances via neighbor-joining phylogeny. We found that in each of the four analyses the Cane Paratore, Mastino Abruzzese, Komondor, and Volpino Italiano were most similar, as well as closest in genomic structure, to the Fomni's Dog. The Sloughi, Anatolian Shepherd, Cane Corso, and Saluki were also identified by STRUCTURE and  $F_{ST}$  as sharing ancestral similarity with the

Fomni's Dog. Like the Fomni's Dog, the Cane Paratore, Mastino Abruzzese, Volpino Italiano, and Cane Corso are breeds with historical origins in Italy. As is suggested by the STRUCTURE output (Figure 4), a common genetic signature is shared across the regional Italian breeds (Cane Corso, Cane Paratore, Fomni's Dog, Mastino Abruzzese, Spinone Italiano, Volpino Italiano, Levriero Meridionale, and Cirneco dell'Etna) together with a group of breeds from the South or East Mediterranean region (Komondor, Saluki, Anatolian Shepherd, Sloughi, and Azawakh). Phylogeny and PCA aid in the resolution of this genetic commonality. PCs distinguish breeds within the STRUCTURE-derived common cluster, isolating the Cane Paratore, Fomni's Dog, Komondor, Levriero Meridionale, Mastino Abruzzese, Spinone Italiano, and Volpino Italiano as distinct from the Cirneco dell'Etna, Cane Corso, Saluki, Sloughi, Azawakh, and Anatolian Shepherd (Figure 6). Phylogenetic inference for the Italian breeds further separates these breeds, identifying a monophyletic clade consisting of Fomni's Dog, Mastino Abruzzese, and Cane Paratore; where each breed still maintains its individuality (Figure 7). These conclusions should be interpreted with caution as a small number of dogs were analyzed for certain breeds. It is unlikely, however, that the addition of more dogs would significantly restructure the PCAs as present separation of breeds is already well defined.

Our analyses also show that the early claims of both Tyndale (1849) and Cetti (1774), who hypothesized that the Fomni's Dog was created through a combination of breeds that resembled sighthounds and molossers, is surprisingly accurate. Instead of a greyhound and a mastiff, however, knowledge of breed origins implicates a Saluki-like coursing hound and a Komondor-like molossoid livestock guardian. Indeed, even the more geographically available sighthound (Pharaoh Hound) and molosser (Neapolitan Mastiff) breeds from Italy are demoted as plausible ancestors through  $F_{ST}$  analysis in favor of the eastern breed-type equivalents. Examination of breed introgression further supports the intertwining of the Fomni's Dog, Komondor, and Saluki.

Our analysis also suggests a directional admixture of Fomni's Dog into an ancestral clade consisting of the Portuguese Water Dog and Cane Paratore. Thus, Fomni's Dog continues to be present in the genetic profiles of other Mediterranean breeds, while other breeds do not appear to have contributed in recent time to the Fomni's Dog. This is unexpected as, given the availability of other island breeds for mating, we would expect that at the genetic level, Fomni's Dog would be little more than a mongrel, which is clearly not the case. Potentially this relates to the prosperity of early human populations on Sardinia, which were at least somewhat dependent on the Fomni's Dog. Thus, the desire to maintain this successful composite breed as it first existed was, and continues to be, strong. By extension, in the absence of strong selection for morphological traits in a breed that retains the critical features of its ancestors, even minor selection of behavioral traits is presumably sufficient to retain homozygosity in genomic regions critical for performance. While such regions still



require fine mapping, our homozygosity analyses demonstrate that such regions exist. A total of 11 genomic regions of putative selection integral for the Foini's Dog have been identified (Table 4). We hypothesize that future fine mapping of these regions will reveal genes necessary for characteristic behaviors of the Foini's Dog.

The association of the Foini's Dog with sighthound and molossoid breeds from the Eastern and Southern Mediterranean, as defined by the present study, is particularly intriguing in light of genetic analyses of modern day humans. Analysis of 300 Italian people against a background of 1272 European, Middle Eastern, and North African individuals revealed that the population of Sardinia shows greatest genetic identity with populations from Hungary, Egypt, Israel, and Jordan (Fiorito *et al.* 2015). The populations from the peninsular regions of Italy, however, share greater similarity with European cultures to the west and north (Fiorito *et al.* 2015). Conversely, human populations across Sardinia show no population stratification based on linguistic or geographic regions, as measured by  $F_{ST}$ , STRUCTURE, or PCA (Di Gaetano *et al.* 2014). This implies that Sardinia was primarily populated by Eastern Mediterranean and North African peoples who, theoretically, brought dogs from their homelands, thus populating the island of Sardinia with the makings of what would in time become the Foini's Dog. Once inhabited, the island maintained a distinctive genetic homogeneity, although geographic and linguistic barriers led to the development of various well-defined subpopulations on the island. The patterns of human population development on Sardinia are thus reflected in the Foini's Dog, primarily when considering its genetic similarity to breeds originating in the Middle East and North Africa (Komondor, Anatolian Shepherd, Sloughi, Saluki, and Azawakh) and its commonalities with other breeds subsequently developed in Italy (Cane Corso, Cane Paratore, Mastino Abruzzese, Spinone Italiano, Volpino Italiano, Levriero Meridionale, and Cirneco dell'Etna).

Through application of GWS and SNP analyses, we have demonstrated that 28 Mediterranean dog breeds present the necessary characteristics of population isolates, the Foini's Dog has obtained equivalent characteristics through distinctive behaviorally-focused selective forces, and the genomic infrastructure of the Foini's Dog mirrors the ancestral human demographic of its vicinity. In addition, the component breeds of the original Foini's Dog are reminiscent of the various peoples that populated the island; highlighting the value of studying other niche populations, particularly in areas of the world where humans and dogs have developed in parallel to garner skills necessary for survival.

## Acknowledgments

The authors thank all the dog owners and their dogs for providing DNA samples. Special thanks to Sir Terence Clark, Mauricio Lima, and Robert Gennari for obtaining DNA samples from regional dog breeds from Italy, North Africa,

and the Middle East. D.L.D., E.A.O, B.W.D, and H.G.P. acknowledge support from the Intramural Program of the International Human Genome Research Institute. This study is part of an Italian study "Detection of morphological, genetic, and behavioral characteristics of Foini's Dog, aimed at selection and official recognition of this breed by ENCI", funded in part by misura 3.13 del Programma Operativo Regione (P. O. R.) Sardegna 2006-2008 (Fondi Europei di Sviluppo Regionale (FESR)).

## Literature Cited

- Alexander, D. H., J. Novembre, and K. Lange, 2009 Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19: 1655–1664.
- American Kennel Club, 2006 *The Complete Dog Book*. Ballantine Books, New York.
- Arason, A., H. Gunnarsson, G. Johannesdottir, K. Jonasson, P. O. Bendahl *et al.*, 2010 Genome-wide search for breast cancer linkage in large Icelandic non-BRCA1/2 families. *Breast Cancer Res.* 12: R50.
- Axelsson, E., A. Ratnakumar, M. L. Arendt, K. Maqbool, M. T. Webster *et al.*, 2013 The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 495: 360–364.
- Baffoe-Bonnie, A. B., J. R. Smith, D. A. Stephan, J. Schleutker, J. D. Carpten *et al.*, 2005 A major locus for hereditary prostate cancer in Finland: localization by linkage disequilibrium of a haplotype in the HPCX region. *Hum. Genet.* 117: 307–316.
- Barbato, M., P. Orozco-terWengel, M. Tapio, and M. W. Bruford, 2015 SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet.* 6: 109.
- Bendjilali, N., W. C. Hsueh, Q. He, D. C. Willcox, C. M. Nievergelt *et al.*, 2014 Who are the Okinawans? Ancestry, genome diversity, and implications for the genetic study of human longevity from a geographically isolated population. *J. Gerontol. A Biol. Sci. Med. Sci.* 69: 1474–1484.
- Benton, M. C., S. Stuart, C. Bellis, D. Macartney-Coxson, D. Eccles *et al.*, 2015 'Mutiny on the Bounty': the genetic history of Norfolk Island reveals extreme gender-biased admixture. *Investig. Genet.* 6: 11.
- Boyko, A. R., 2011 The domestic dog: man's best friend in the genomic era. *Genome Biol.* 12: 216.
- Bresciani, A., 1850 *Dei costumi dell'isola di Sardegna comparati cogli antichissimi popoli orientali: per Antonio Bresciani*. Uffizio della Civiltà Cattolica, Naples, Italy.
- Capocasa, M., P. Anagnostou, V. Bachis, C. Battaglia, S. Bertoni *et al.*, 2014 Linguistic, geographic and genetic isolation: a collaborative study of Italian populations. *J. Anthropol. Sci.* 92: 201–231.
- Cetti, F., 1774 *I Quadrupedi di Sardegna*. Kessinger Publishing, LLC.
- Chiang, A. P., D. Nishimura, C. Searby, K. Elbedour, R. Carmi *et al.*, 2004 Comparative genomic analysis identifies an ADP-ribosylation factor-like gene as the cause of Bardet-Biedl syndrome (BBS3). *Am. J. Hum. Genet.* 75: 475–484.
- Cingolani, P., and A. Platts, L. le Wang, M. Coon, T. Nguyen *et al.*, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 6: 80–92.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks *et al.*, 2011 The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

- DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire *et al.*, 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43: 491–498.
- Di Gaetano, C., G. Fiorito, M. F. Ortu, F. Rosa, S. Guarrera *et al.*, 2014 Sardinians genetic background explained by runs of homozygosity and genomic regions under positive selection. *PLoS One* 9: e91237.
- Edwardes, C., 1889 *Sardinia and the Sardes*. Richard Bentley and Son, London.
- Farag, T. I., and A. S. Teebi, 1989 High incidence of Bardet Biedl syndrome among the Bedouin. *Clin. Genet.* 36: 463–464.
- Felsenstein, J., 1989 PHYLIP: Phylogeny Inference Package. *Cladistics* 5: 164–166.
- Fiorito, G., C. Di Gaetano, S. Guarrera, F. Rosa, M. W. Feldman *et al.*, 2015 The Italian genome reflects the history of Europe and the Mediterranean basin. *Eur. J. Hum. Genet.* 24: 1056–1062.
- Fogel, B., 1995 *The Encyclopedia of the Dog*. Dorling Kindersley, New York.
- Francisco, L. V., A. A. Langston, C. S. Mellers, C. L. Neal, and E. A. Ostrander, 1996 A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mamm. Genome* 7: 359–362.
- Galibert, F., P. Quignon, C. Hitte, and C. Andre, 2011 Toward understanding dog evolutionary and domestication history. *C. R. Biol.* 334: 190–196.
- Georgi, B., D. Craig, R. L. Kember, W. Liu, I. Lindquist *et al.*, 2014 Genomic view of bipolar disorder revealed by whole genome sequencing in a genetic isolate. *PLoS Genet.* 10: e1004229.
- Ginns, E. I., P. St Jean, R. A. Philibert, M. Galdzicka, P. Damschroder-Williams *et al.*, 1998 A genome-wide search for chromosomal loci linked to mental health wellness in relatives at high risk for bipolar affective disorder among the Old Order Amish. *Proc. Natl. Acad. Sci. USA* 95: 15531–15536.
- Haider, N. B., R. Carmi, H. Shalev, V. C. Sheffield, and D. Landau, 1998 A Bedouin kindred with infantile nephronophthisis demonstrates linkage to chromosome 9 by homozygosity mapping. *Am. J. Hum. Genet.* 63: 1404–1410.
- Hartikainen, J. M., H. Tuhkanen, V. Kataja, A. M. Dunning, A. Antoniou *et al.*, 2005 An autosome-wide scan for linkage disequilibrium-based association in sporadic breast cancer cases in eastern Finland: three candidate regions found. *Cancer Epidemiol. Biomarkers Prev.* 14: 75–80.
- Hayward, J. J., M. G. Castelhana, K. C. Oliveira, E. Corey, C. Balkman *et al.*, 2016 Complex disease and phenotype mapping in the domestic dog. *Nat. Commun.* 7: 10460.
- Hicks, A. A., H. Petursson, T. Jonsson, H. Stefansson, H. S. Johannsdottir *et al.*, 2002 A susceptibility gene for late-onset idiopathic Parkinson's disease. *Ann. Neurol.* 52: 549–555.
- Hou, L., G. Faraci, D. T. Chen, L. Kassem, T. G. Schulze *et al.*, 2013 Amish revisited: next-generation sequencing studies of psychiatric disorders among the Plain people. *Trends Genet.* 29: 412–418.
- Hsueh, W. C., P. L. St Jean, B. D. Mitchell, T. I. Pollin, W. C. Knowler *et al.*, 2003 Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21-q24. *Diabetes* 52: 550–557.
- Kainu, T., S. H. Juo, R. Desper, A. A. Schaffer, E. Gillanders *et al.*, 2000 Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc. Natl. Acad. Sci. USA* 97: 9603–9608.
- Karason, A., J. E. Gudjonsson, H. H. Jonsson, V. B. Hauksson, E. H. Runarsdottir *et al.*, 2005 Genetics of psoriasis in Iceland: evidence for linkage of subphenotypes to distinct Loci. *J. Invest. Dermatol.* 124: 1177–1185.
- Kember, R. L., B. Georgi, J. E. Bailey-Wilson, D. Stambolian, S. M. Paul *et al.*, 2015 Copy number variants encompassing Mendelian disease genes in a large multigenerational family segregating bipolar disorder. *BMC Genet.* 16: 27.
- Korneliusson, T. S., A. Albrechtsen, and R. Nielsen, 2014 ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* 15: 356.
- Leroy, G., E. Verrier, J. C. Meriaux, and X. Rognon, 2009 Genetic diversity of dog breeds: between-breed diversity, breed assignment and conservation approaches. *Anim. Genet.* 40: 333–343.
- Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Lindblad-Toh, K., C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe *et al.*, 2005 Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438: 803–819.
- Marsden, C. D., D. Ortega-Del Vecchyo, D. P. O'Brien, J. F. Taylor, O. Ramirez *et al.*, 2016 Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proc. Natl. Acad. Sci. USA* 113: 152–157.
- Mellanby, R. J., R. Ogden, D. N. Clements, A. T. French, A. G. Gow *et al.*, 2013 Population structure and genetic heterogeneity in popular dog breeds in the UK. *Vet. J.* 196: 92–97.
- Nyström-Lahti, M., P. Sistonen, J. P. Mecklin, L. Pylkkanen, L. A. Aaltonen *et al.*, 1994 Close linkage to chromosome 3p and conservation of ancestral founding haplotype in hereditary non-polyposis colorectal cancer families. *Proc. Natl. Acad. Sci. USA* 91: 6054–6058.
- Ostrander, E. A., G. F. Sprague, Jr., and J. Rine, 1993 Identification and characterization of dinucleotide repeat (CA)<sub>n</sub> markers for genetic mapping in dog. *Genomics* 16: 207–213.
- Parker, H. G., L. V. Kim, N. B. Sutter, S. Carlson, T. D. Lorentzen *et al.*, 2004 Genetic structure of the purebred domestic dog. *Science* 304: 1160–1164.
- Parra, D., S. Mendez, J. Canon, and S. Dunner, 2008 Genetic differentiation in pointing dog breeds inferred from microsatellites and mitochondrial DNA sequence. *Anim. Genet.* 39: 1–7.
- Patterson, N., A. L. Price, and D. Reich, 2006 Population structure and eigenanalysis. *PLoS Genet.* 2: e190.
- Peltonen, L., A. Palotie, and K. Lange, 2000 Use of population isolates for mapping complex traits. *Nat. Rev. Genet.* 1: 182–190.
- Pickrell, J. K., and J. K. Pritchard, 2012 Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* 8: e1002967.
- Pickrell, J. K., N. Patterson, C. Barbieri, F. Berthold, L. Gerlach *et al.*, 2012 The genetic prehistory of southern Africa. *Nat. Commun.* 3: 1143.
- Pribanova, M., P. Horak, D. Schroffelova, T. Urban, R. Bechynova *et al.*, 2009 Analysis of genetic variability in the Czech Dachshund population using microsatellite markers. *J. Anim. Breed. Genet.* 126: 311–318.
- Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick *et al.*, 2006 Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38: 904–909.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira *et al.*, 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81: 559–575.
- Raj, A., M. Stephens, and J. K. Pritchard, 2014 fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197: 573–589.

- Rambaut, A., 2014 FigTree, version July 2014. Edinburgh. <http://http://tree.bio.ed.ac.uk/software/figtree/>.
- Rimbault, M., and E. A. Ostrander, 2012 So many doggone traits: mapping genetics of multiple phenotypes in the domestic dog. *Hum. Mol. Genet.* 21: R52–R57.
- Rokman, A., A. B. Baffoe-Bonnie, E. Gillanders, H. Fredriksson, V. Autio *et al.*, 2005 Hereditary prostate cancer in Finland: fine-mapping validates 3p26 as a major predisposition locus. *Hum. Genet.* 116: 43–50.
- Rund, D., T. Cohen, D. Filon, C. E. Dowling, T. C. Warren *et al.*, 1991 Evolution of a genetic disease in an ethnic isolate: beta-thalassemia in the Jews of Kurdistan. *Proc. Natl. Acad. Sci. USA* 88: 310–314.
- Sajantila, A., A. H. Salem, P. Savolainen, K. Bauer, C. Gierig *et al.*, 1996 Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. *Proc. Natl. Acad. Sci. USA* 93: 12035–12039.
- Sambrook, J., E. F. Fritsch, and T. Maniatis, 1989 *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sarantaus, L., P. Huusko, H. Eerola, V. Launonen, P. Vehmanen *et al.*, 2000 Multiple founder effects and geographical clustering of BRCA1 and BRCA2 families in Finland. *Eur. J. Hum. Genet.* 8: 757–763.
- Schoenebeck, J. J., and E. A. Ostrander, 2014 Insights into morphology and disease from the dog genome project. *Annu. Rev. Cell Dev. Biol.* 30: 535–560.
- Scott, D. A., R. Carmi, K. Elbedour, S. Yosefsberg, E. M. Stone *et al.*, 1996 An autosomal recessive nonsyndromic-hearing-loss locus identified by DNA pooling using two inbred Bedouin kindreds. *Am. J. Hum. Genet.* 59: 385–391.
- Seboun, E., A. Lemainque, and C. E. Jackson, 2005 Amish brittle hair syndrome gene maps to 7p14.1. *Am. J. Med. Genet. A.* 134: 290–294.
- Sechi, S., M. Polli, S. Marelli, A. Talenti, P. Crepaldi *et al.*, 2016 Fonnì's Dog: morphological and genetic characteristics for a breed standard definition. *Ital. J. Anim. Sci.* (in press).
- Sheffield, V. C., R. Carmi, A. Kwitek-Black, T. Rokhlina, D. Nishimura *et al.*, 1994 Identification of a Bardet-Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. *Hum. Mol. Genet.* 3: 1331–1335.
- Sidore, C., F. Busonero, A. Maschio, E. Porcu, S. Naitza *et al.*, 2015 Genome sequencing elucidates Sardinian genetic architecture and augments association analyses for lipid and blood inflammatory markers. *Nat. Genet.* 47: 1272–1281.
- Simpson, M. A., H. E. Cross, L. Cross, M. Helmuth, and A. H. Crosby, 2009 Lethal cystic kidney disease in Amish neonates associated with homozygous nonsense mutation of NPHP3. *Am. J. Kidney Dis.* 53: 790–795.
- Stone, D. L., R. Agarwala, A. A. Schaffer, J. L. Weber, D. Vaske *et al.*, 1998 Genetic and physical mapping of the McKusick-Kaufman syndrome. *Hum. Mol. Genet.* 7: 475–481.
- Suarez, N. M., E. Betancor, R. Fregel, and J. Pestano, 2013 Genetic characterization, at the mitochondrial and nuclear DNA levels, of five Canary Island dog breeds. *Anim. Genet.* 44: 432–441.
- Szpiech, Z. A., J. Xu, T. J. Pemberton, W. Peng, S. Zollner *et al.*, 2013 Long runs of homozygosity are enriched for deleterious variation. *Am. J. Hum. Genet.* 93: 90–102.
- Thorgeirsson, T. E., H. Oskarsson, N. Desnica, J. P. Kostic, J. G. Stefansson *et al.*, 2003 Anxiety with panic disorder linked to chromosome 9q in Iceland. *Am. J. Hum. Genet.* 72: 1221–1230.
- Tyndale, J. W., 1849 *The Island of Sardinia, including pictures of the manners and customs of the Sardinians, and notes on the antiquities and modern objects of interest in the island: to which is added some account of the House of Savoy*. Richard Bentley, New Burlington Street, London.
- Van der Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel *et al.*, 2013 From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43: 11.10.11–11.10.33.
- Wayne, R. K., and B. M. vonHoldt, 2012 Evolutionary genomics of dog domestication. *Mamm. Genome* 23: 3–18.
- Zhai, G., J. Zhou, M. O. Woods, J. S. Green, P. Parfrey *et al.*, 2016 Genetic structure of the Newfoundland and Labrador population: founder effects modulate variability. *Eur. J. Hum. Genet.* 24: 1063–1070.

Communicating editor: R. Nielsen