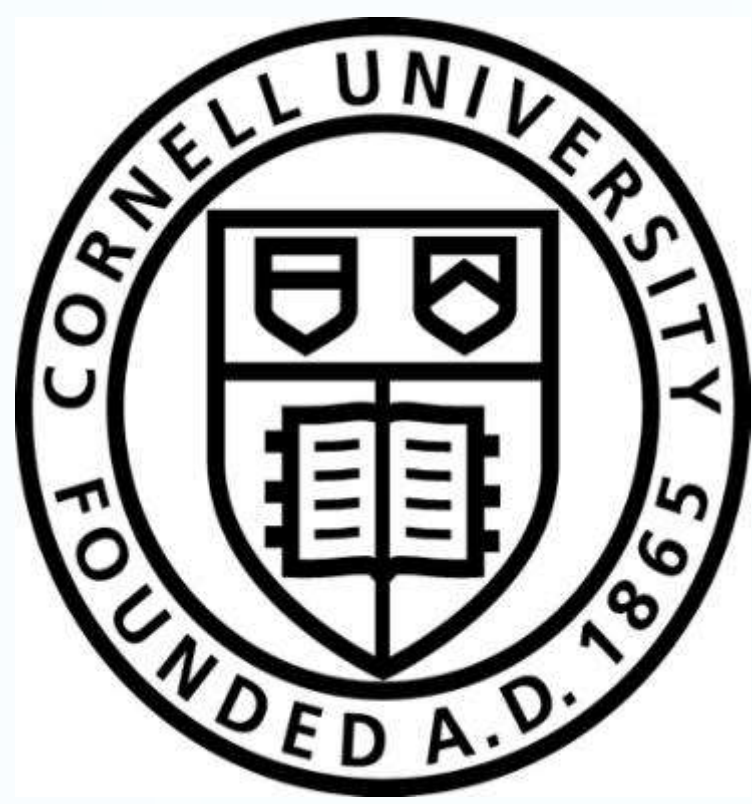


Signals of Selection in Alaskan Sled Dogs



Alexander J. Valenti¹, Adam Boyko, PhD², Heather J. Huson, PhD¹

¹Cornell University College of Agriculture and Life Science, Department of Animal Science

²Cornell University, College of Veterinary Medicine, Biomedical Sciences

Background and Objectives

Modern racing sled dogs are a good illustration of how working dogs are selected and how a breed is created. Early sled dog races were recreational tests of the prowess of hard-working, freight-pulling dogs and of the drivers tasked with transporting cargo across the snowy frontiers of North America and Asia. These tests rapidly evolved into sprint and distance races, consequently creating two classifications of dogs, sprinters and distance runners. Sprinters have been optimized for speed, traversing courses of up to 30 miles. Distance runners have undergone selection for extreme endurance, running races covering over 1000 miles. Multiple breeds have been introduced into the sled-dog population in an attempt to enhance performance. Today's best racing dogs are the product a breeding and selection process based on their athletic ability combined with the proper behaviors necessary for a team to function as a single unit. Using several genome analysis tests, mainly focusing on runs of homozygosity, this study aims to analyze the impact of this selection process and map evolutionary relationships between the two classifications of sled dogs as well as a comparison to a select group of pure bred dogs.

SPRINT



DISTANCE



	SPRINT	DISTANCE
Average speed	18-25 mph	8-12 mph
Average distance	10-30 miles	900 – 1000 miles
Average sled weight	12 - 15 lbs	30 - 40 lbs
Average number of dogs per team	8 – 18	16 - 20

Methods

- Sample Set: 325 total dogs
 - 158 Alaskan Sled dogs
 - 167 Pure bred dogs from breeds previously identified as hybridized with the Alaskan sled dog based on genetic research and pedigree information

Breed/Classification	# of Individuals
Sprint	93
Distance	65
Alaskan Malamute	12
Border Collie	67
Borzo	12
German Shorthaired	11
German Wired	2
Greyhound	22
Italian Greyhound	6
Saluki	7
Siberian Husky	17
Village-Alaska	11

- Genotypic Information
 - Blood samples drawn and analyzed using the Illumina Canine HD Beadchip and whole genome sequencing
- Golden Helix (PCA, ROH, Fst, GWAS)
 - Quality Control Thresholds
 - Drop if Call Rate < .9
 - Drop if MAF < .05
 - Quality Control Results
 - 113,113 SNPs remained of 113,172 following QC
- PANTHER Gene Pathway Analysis

Results

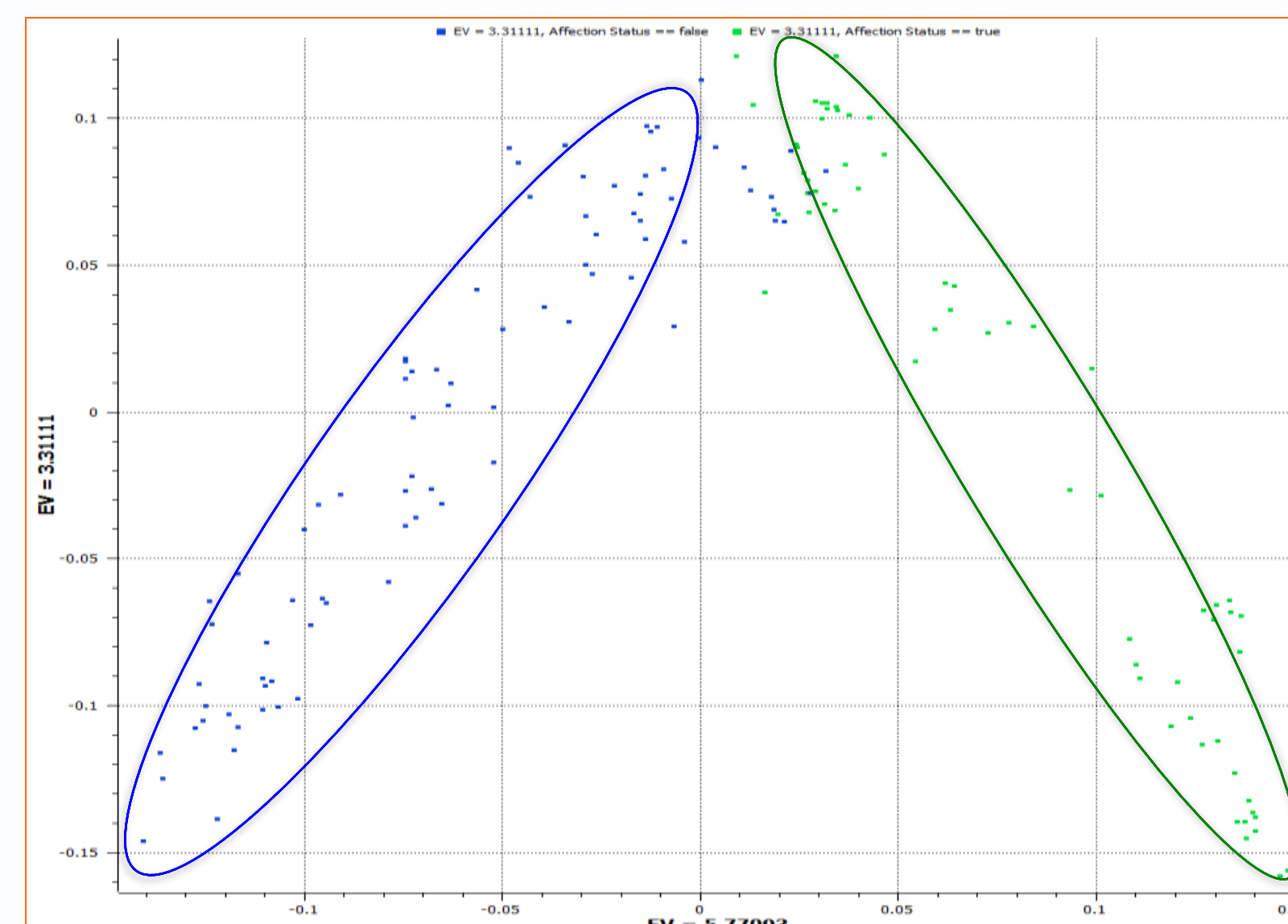


Figure 1: Principal Component Analysis identifies population structure within Alaskan Sled dogs; PC 1 on the x-axis, PC 2 on the y-axis

- PCA indicating genetic differentiation of sprint (blue) and distance (green) dogs
- PCA used to correct for population structure in the ROH association analyses

- Principal Component Analysis (Fig 1)
 - Identifies population structure
- Fixation Index (Fst) by Marker (Fig 2 and 3)
 - Indicated genetic divergence by marker
- ROH Analysis
 - Runs were grouped by size for both populations
 - Average total number of ROH was calculated for each population and also analyzed by length of run (See Fig 4, Fig 5, Fig 6, and Fig 7)
 - Thresholds
 - Distance set to 500 kB with 15 as the minimum #SNPs
 - Allowed runs to contain up to 3 missing genotypes with the max gap between SNPs in a run set to 200 kB
 - Numeric association ran between all subsets of populations using PCA to correct for population structure (See Fig 8 and 9)
- PANTHER Gene Pathway Analysis using statistically significant runs of Sprint and Distance animals

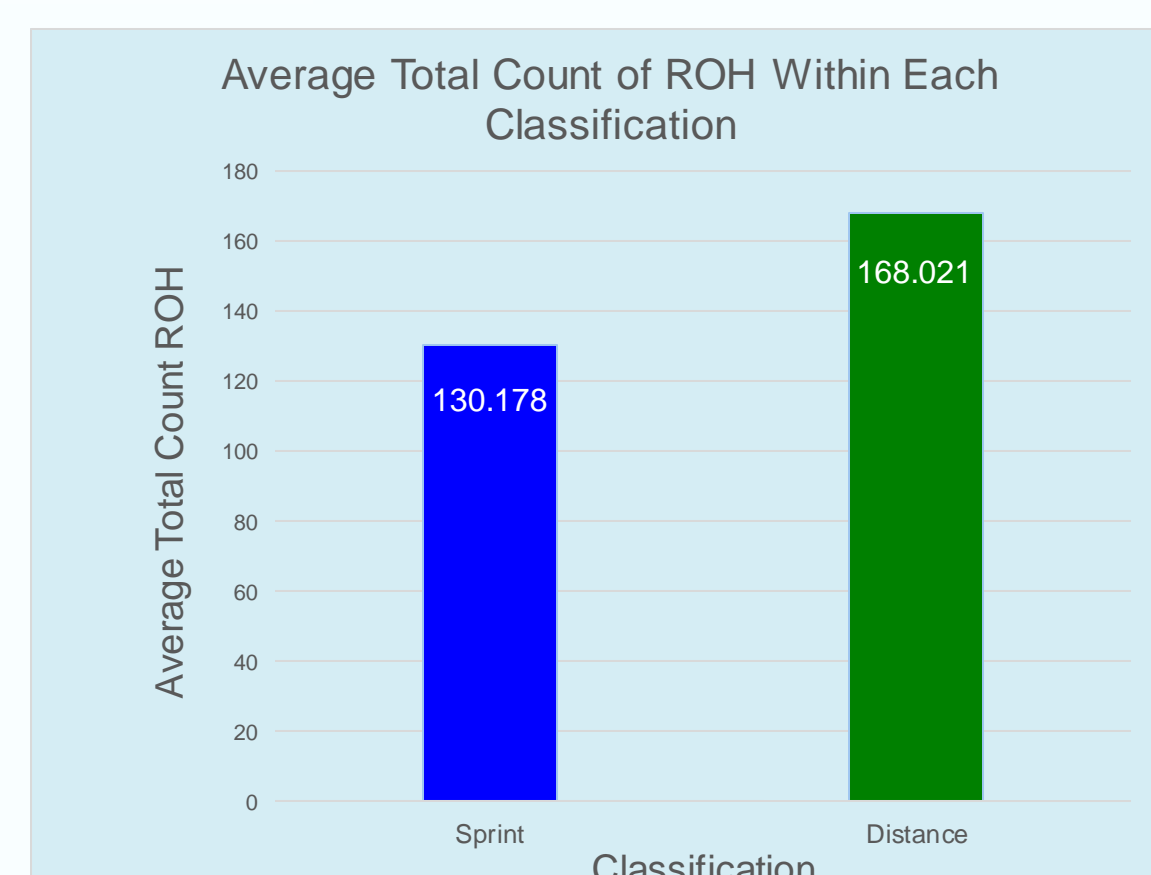


Figure 4: Average total count of ROH per individual between Sprint and Distance animals

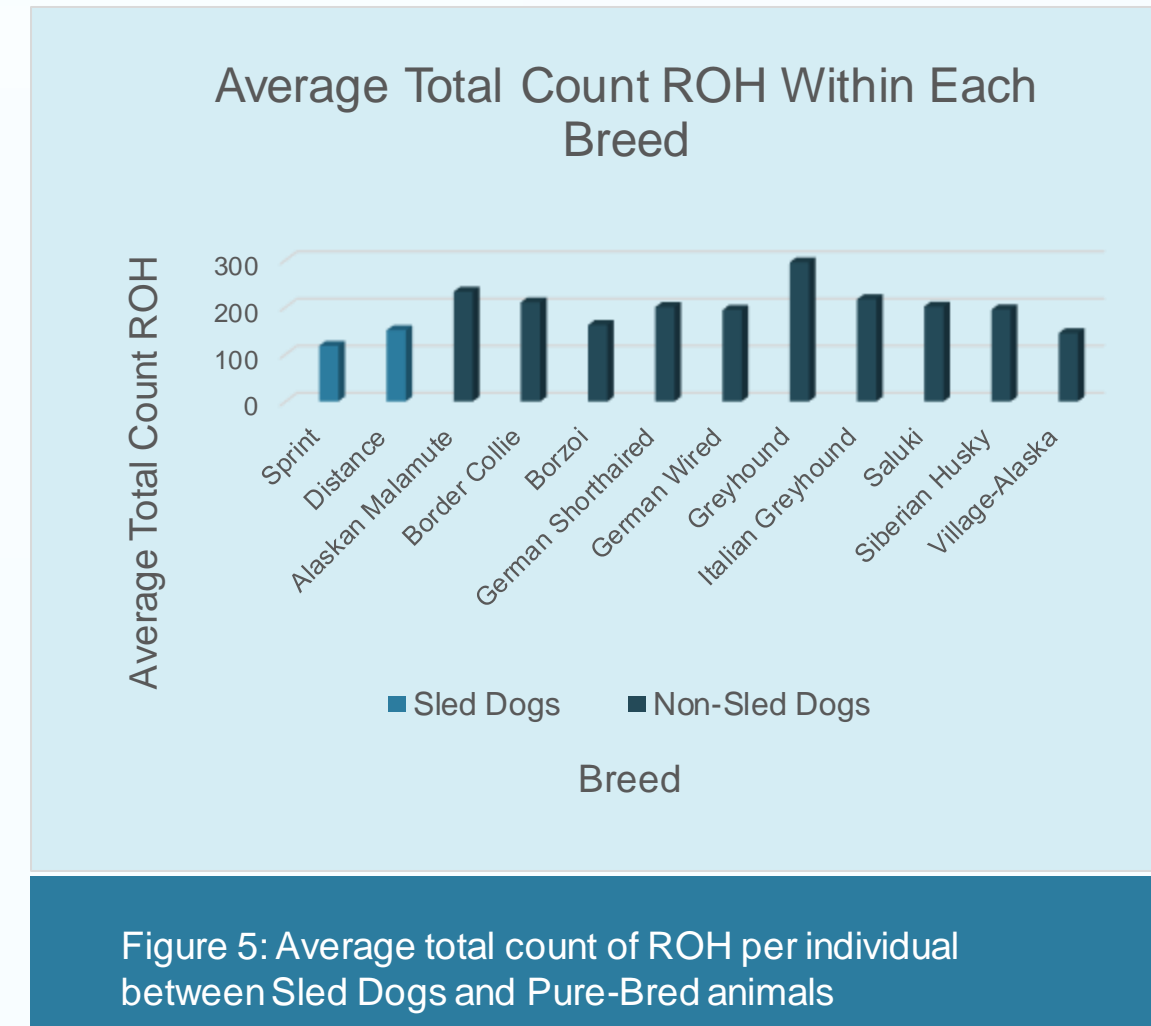


Figure 5: Average total count of ROH per individual between Sled Dogs and Pure-Bred animals

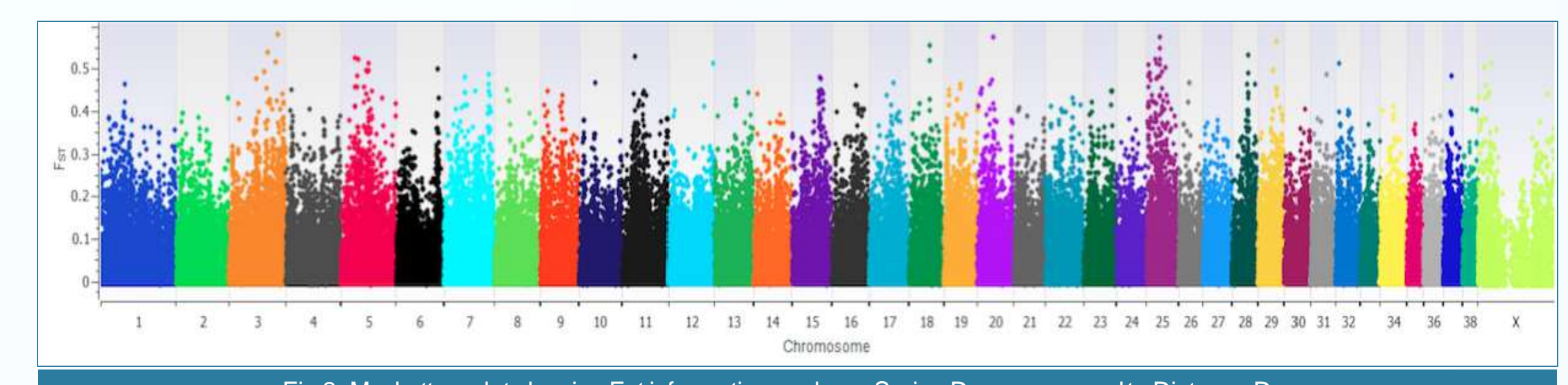


Figure 2: Manhattan plot showing Fst informative markers: Spring Dogs compared to Distance Dogs

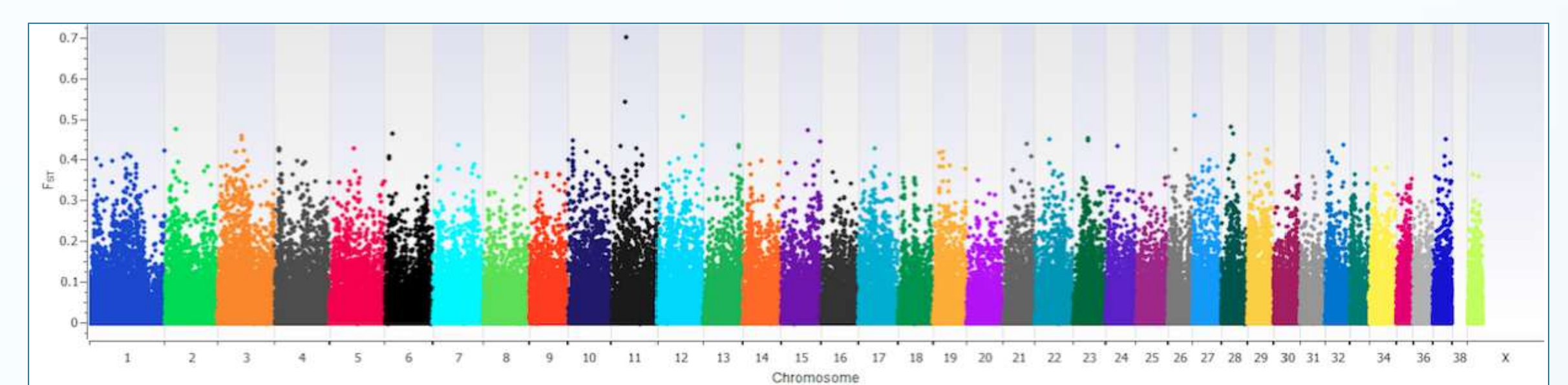


Figure 3: Manhattan plot showing Fst informative markers: Sled Dogs compared to Pure-Bred Dogs

ROH Association Analysis: Clusters of ROH were identified in the respective sample populations using Golden Helix SVS software. An ROH cluster is represented by the first SNP within that cluster. Each ROH cluster was analyzed for association to either Alaskan sled dog racing classification (Fig 8) or comparing Alaskan sled dogs to pure bred breeds (Fig 9).

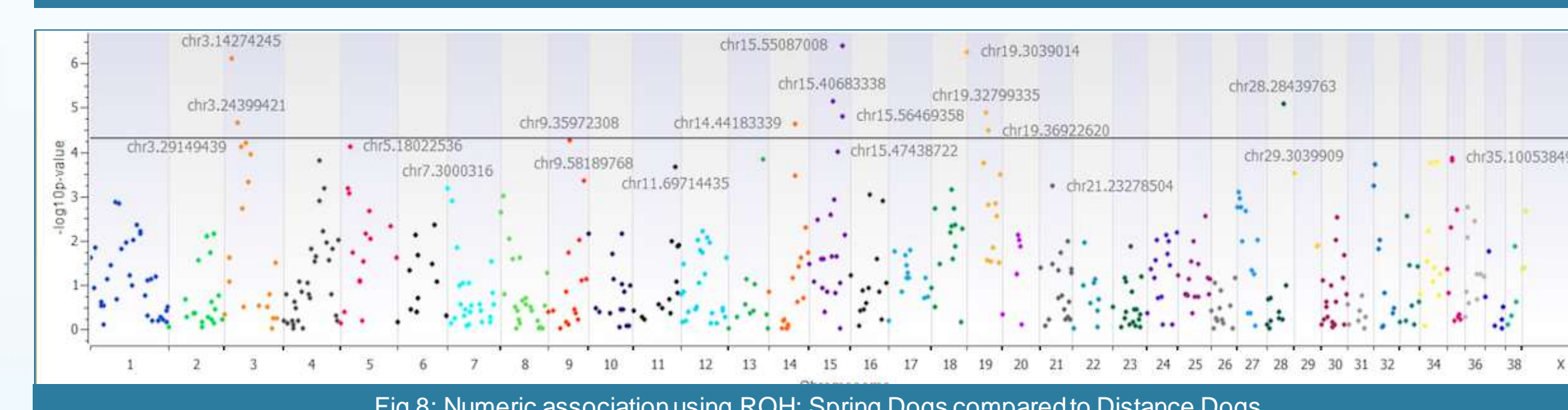


Figure 8: Numeric association using ROH: Spring Dogs compared to Distance Dogs

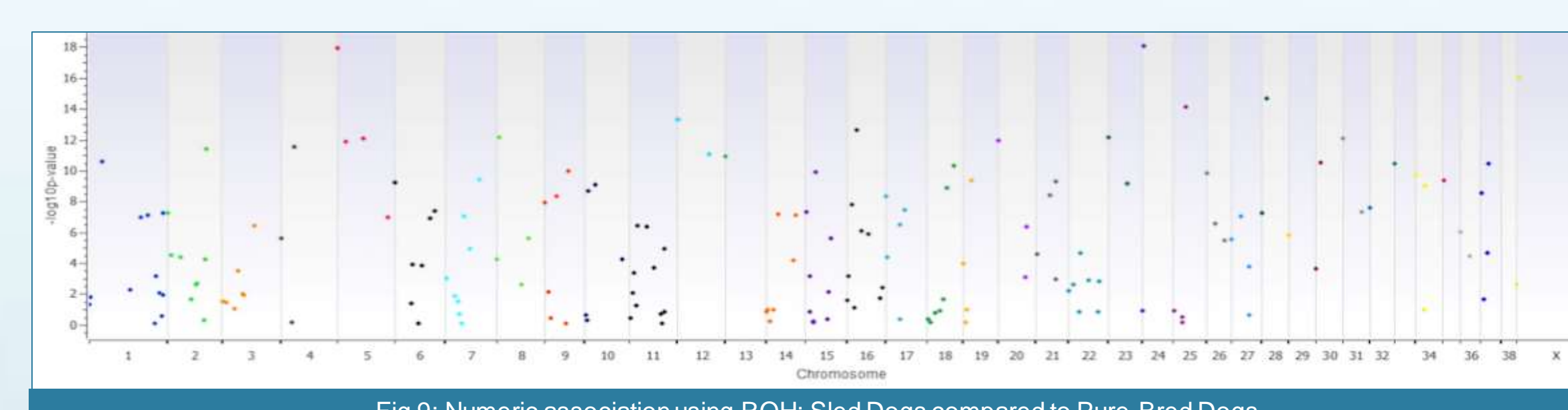


Figure 9: Numeric association using ROH: Sled Dogs compared to Pure-Bred Dogs

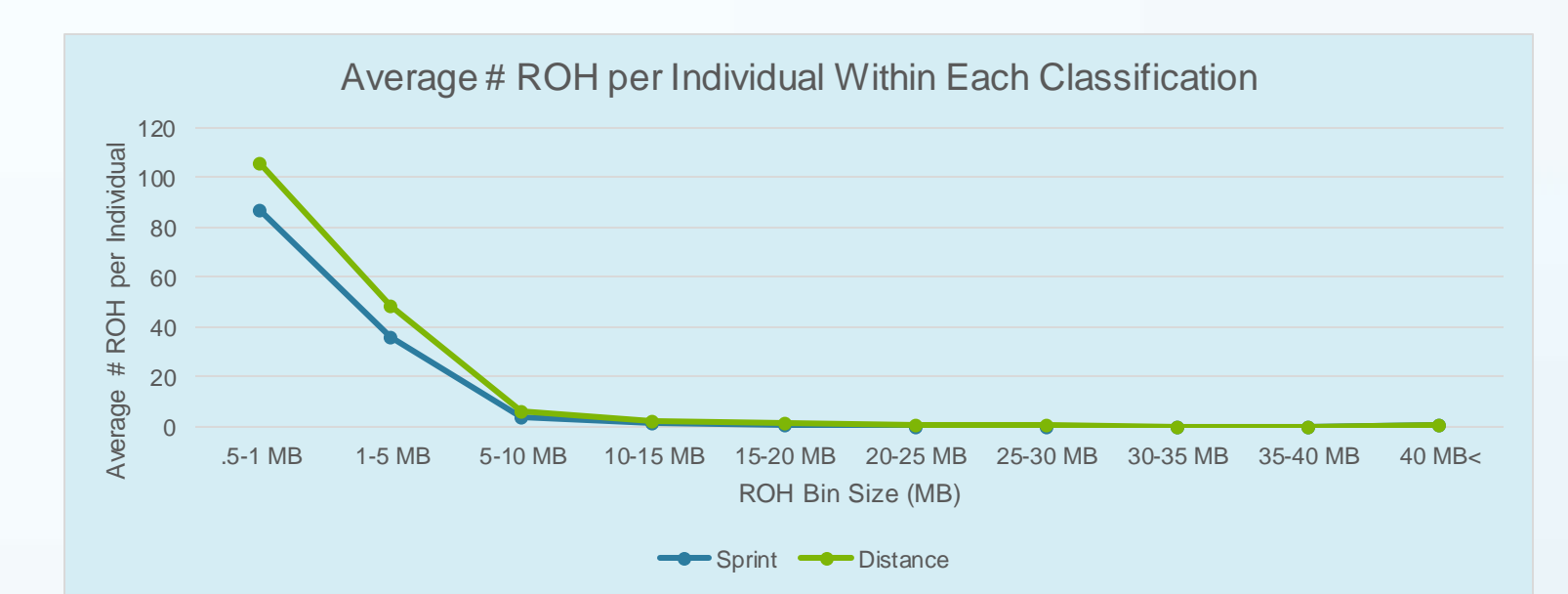


Figure 6: Average #ROH per individual by length of run: Sprint Dogs vs Distance Dogs

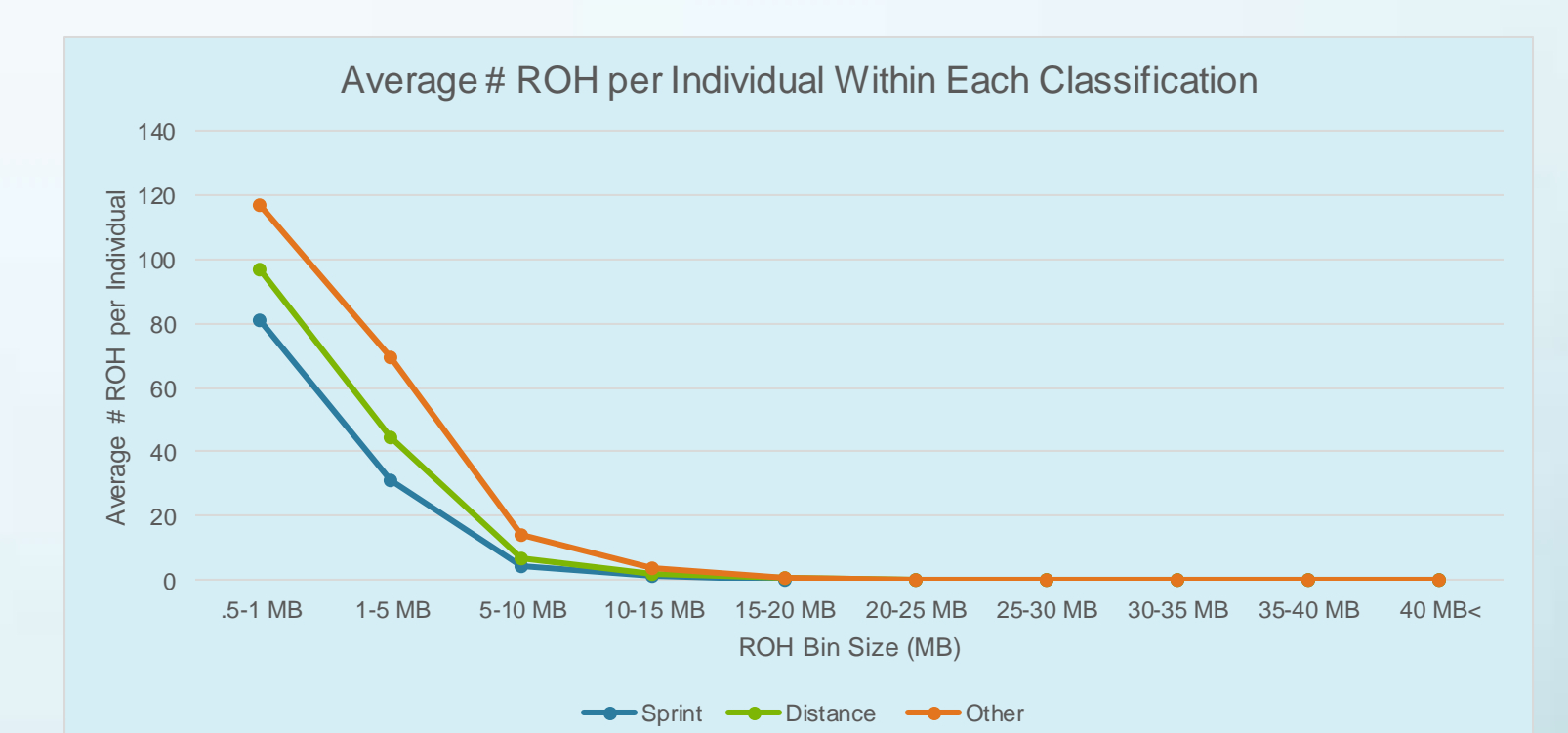


Figure 7: Average #ROH per individual by length of run: Sled Dogs vs Pure-Bred Dogs

Discussion and Future Direction

The numeric association tests identified statistically significant runs of homozygosity that were analyzed using UCSC Genome Browser to search for possible candidate genes. Once classified, these candidate genes were imputed into the gene ontology program PANTHER in order to distinguish any relevant biological pathways associated with the ROH regions in the subsequent populations.

- PCA identified population structure within the Alaskan Sled dogs segregating dogs based on classification.
- F_{ST} identified population informative markers between classification of sled dogs and between purebred and Alaskan sled dogs indicating selection.
 - Sprint Dogs vs Distance Dogs: Several large peaks on chromosomes 1, 18, 20, 28, and 29 identify areas of interest regarding genetic divergence
 - Alaskan Sled Dogs vs Pure-Bred Dogs: One large peak on chromosome 11 indicates a strong possibility of genetic divergence in that region
- Average ROH in terms of the count and length were identified and suggest increased inbreeding within distance sled dogs as compared to sprint sled dogs. In general, sled dogs have fewer ROH than purebred dogs.
- Statistically significant ROH association between Sprint and Distance animals have found several genes with biological implications to
 - locomotion, blood circulation, muscle contraction and development, skeletal and nervous system development, and response to external stimuli.
- Moving forward
 - identify and categorize relevant genes and associated biological pathways using the numeric association between runs in the sled dog population vs. runs in the pure-bred population.

