Successes and challenges of genome-wide association (GWA) analyses in cats

Hasan Alhaddad[1], B. Gandolfi[2], and L. A. Lyons[2]

1) Department of Biological Science, Kuwait University, Kuwait
2) College of Veterinary Medicine, University of Missouri-Columbia, Columbia, MO

Introduction

- Genome sequencing projects enabled the identification of millions of single nucleotide polymorphisms (SNPs) between individuals. Genotyping SNPs in large number of individuals using the high throughput technology and using population based analyses allowed identifying disease and phenotype mutation without the need of an extended pedigree.
- The sequencing of several cat genomes resulted in the development of **illumina infinium iSelect 63K cat DNA array**.
- The 63K Feline array are contains genome-wide markers that can be genotyped at a low cost and assist in the identification of cat diseases and aesthetic traits.

Objective

Test the utility and power of the 63K Feline SNP chip in performing genome-wide association studies (GWAS) for autosomal recessive, autosomal dominant, and X-linked phenotypic traits.

Dataset and Analyses

- A comprehensive genotype dataset of over 2000 cats (breeds and random bred) was used as the sample source of the analyses. For each GWAS, a subset of cases and control were carefully selected (**Fig. 1**).
- Case-control GWA analyses were performed using the open source program *PLINK*.
- For three autosomal recessive traits (dilute coloration, point coloration, and long hair), the causative variant is on the array and was genotyped. Linkage disequilibrium (LD) between the causative marker and adjacent markers was calculated using the *genetics* package in R, which allowed estimating the power of the array [6].

1. Autosomal recessive trait in random bred population

- GWA analysis of autosomal recessive **dilute color** [1] using random bred cases and controls resulted in a single significant marker (**Fig. 1a**). The significant marker is the causative marker intentionally placed on the array.
- The absence of other associated markers is due to the low LD between the causative marker and nearby SNPs (**Fig. 1a-inner plot)**. For the closest marker to have similar association power, the number of samples need to be increased from 114 to 427.

2. Autosomal recessive trait in a breed and under selection

- GWA analysis of autosomal recessive **point color** [2] using Himalayan (pointed Persian) cats as cases and Persian cats as controls resulted in a large number of significant markers (**Fig. 1b**).
- The presence of many associated markers in addition to the causative mutation is due to the high LD between the causative marker and near by and distant SNPs (**Fig. 1b-inner plot)**. The high number of linked markers to the causative mutation is a result of the artifical selection. For the closest marker to have similar association power, the number of samples need to be increased from 49 to 50.

3. Autosomal recessive trait in a breed without selection

- GWA analysis of autosomal recessive **long hair** [3] using La Perm cats cases and controls resulted in several significantly associated markers (**Fig. 1c**).
- Adjacent markers are in high LD with the causative mutation (**Fig. 1c-inner plot)**. Due to the absence of intense selection pressure for the trait in the breed, the number of linked markers is smaller than the point color example (above). For the closest marker to have similar association power, the number of samples need to be increased from 54 to 66.

4. Autosomal dominant trait in a breed and under selection

- GWA analysis of autosomal dominant **curly hair** [4] using Selkirk Rex cats as cases and straight hair Selkirk Rex and Persian cats as controls resulted in a number of significant markers (**Fig. 1d**). The association analysis resulted in the identification of the causative mutation as published in [4].
- The dominant trait was identified with as low as 9 cases and 29 controls. The successful association outcome is likely due to the selection pressure on the trait.

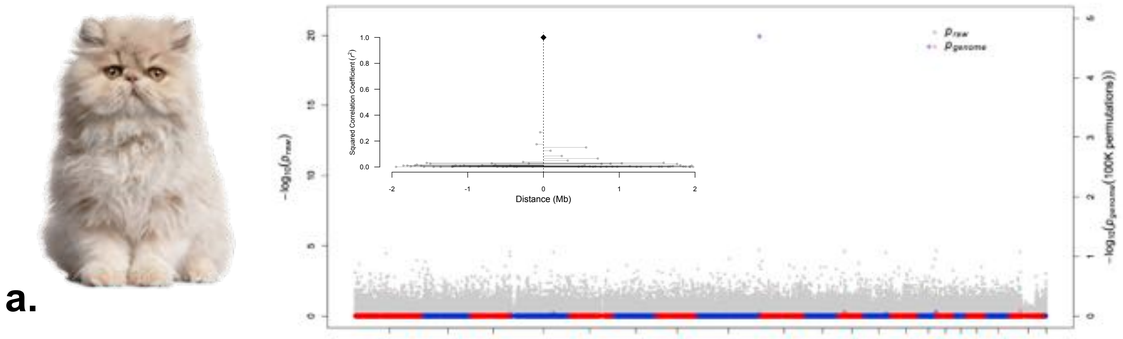
5. X-linked trait in random bred population

- GWA analysis of X-linked **orange color** [5] using random bred cases and controls resulted in a few significant markers (**Fig. 1e**). The associated markers reside in the same linkage region previously identified [5].
- The placement of the orange color locus in the same published region using as little as 24 cases suggests a power in the current density of the array for detecting association on the X-chromosome.

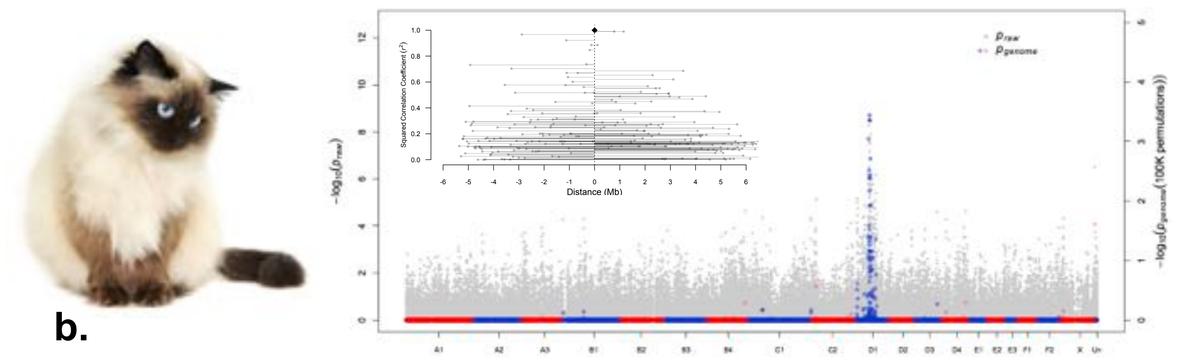
Conclusion

- **The current density of the Feline array (63K) is sufficient to detect association of recessive and dominant traits/diseases (1) under selection, (2) in small populations, and (3) resulting from a recent mutation.**
- **X-chromosome markers are likely to detect association for sex specific traits/ diseases.**
- **The current density of the array is not enough to detect association in random bred populations.**

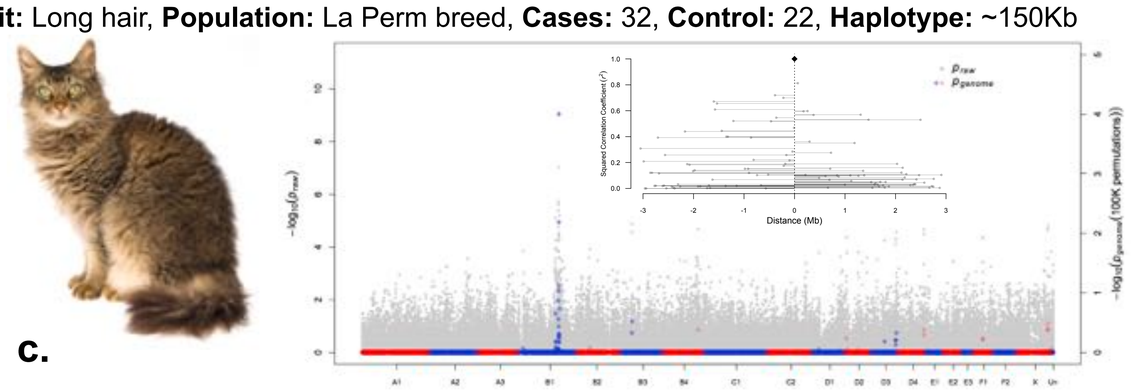
Trait: Dilute color, **Population:** Random bred, **Cases:** 33, **Control:** 81, **Haplotype:** NA



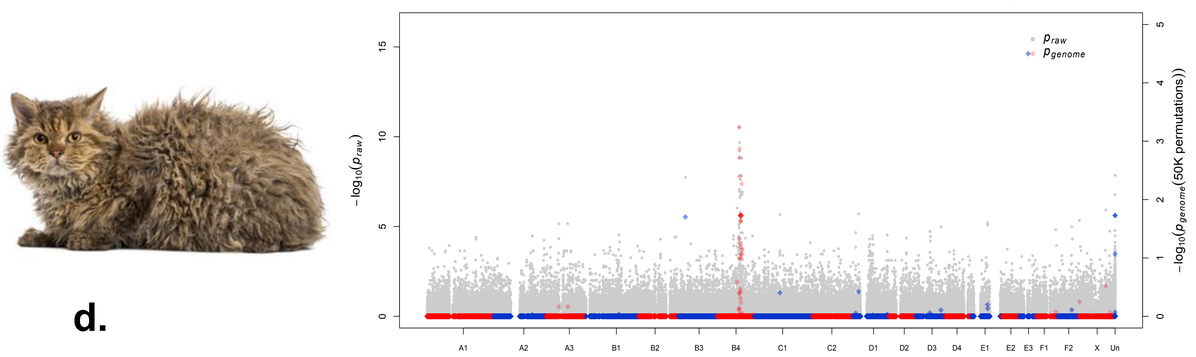
Trait: Point color, **Population:** Persian breed, **Cases:** 21, **Control:** 28, **Haplotype:** ~1Mb



Trait: Long hair, **Population:** La Perm breed, **Cases:** 32, **Control:** 22, **Haplotype:** ~150Kb



Trait: Curly hair, **Population:** Selkirk Rex breed, **Cases:** 9, **Control:** 29, **Haplotype:** ~600Kb



Trait: Orange color, **Population:** Random bred, **Cases:** 24, **Control:** 69, **Haplotype:** ~1.5Mb

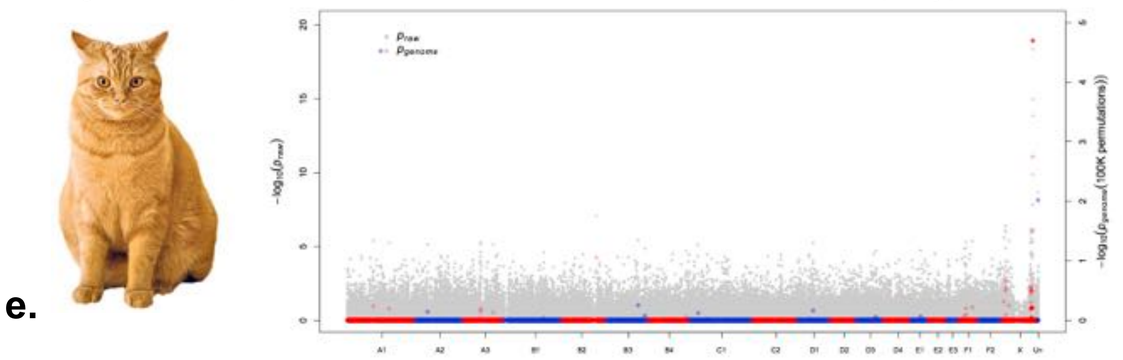


Fig.1. Illustrative genome-wide association analyses for five phenotypic traits in the domestic cats. a-c) Remapping of three autosomal recessive traits using different populations. Causative variants of the three traits were previously identified [1-3] and markers are included on the 63K SNP chip. **d)** Reproducing GWAS of an autosomal dominant trait [4]. **e)** GWAS of X-linked trait that was previously localized to X chromosome region [5]. Manhattan plots (a-e) of the association analyses where x-axis represent chromosomes, gray dots and left y-axis represent raw P-values of the association, and red/blue dots and right y-axis represent the permuted P-values. Inner graphs in (a-c) shows the LD between the trait's causative mutation (black dot) and adjacent markers (gray dots).

References

1 Ishida Y, David VA, Eizirik E, Schaffer AA, Neelam BA, Roelke ME, et al. (2006) A homozygous single-base deletion in MLPH causes the dilute coat color phenotype in the domestic cat. Genomics 88: 698-705.

2 Lyons LA, Imes DL, Rah HC, Grahn RA (2005) Tyrosinase mutations associated with Siamese and Burmese patterns in the domestic cat (Felis catus). Anim Genet 36: 119-126.

3 Drogemuller C, Rufenacht S, Wichert B, Leeb T (2007) Mutations within the *FGF5* gene are associated with hair length in cats. Anim Genet 38: 218-221.

4 Gandolfi B, Alhaddad H, Joslin SE, Khan R, Filler S, Brem G, et al. (2013) A splice variant in KRT71 is associated with curly coat phenotype of Selkirk Rex cats. Sci Rep 3: 2000.

5 Grahn RA, Lemesch BM, Millon LV, Matise T, Rogers QR, Morris JG, et al. (2005) Localizing the X-linked orange colour phenotype using feline resource families. Anim Genet 36: 67-70.

6 Pritchard JK, Przeworski M (2001) Linkage disequilibrium in humans: models and data. Am J Hum Genet 69: 1-14.