AD_____

Award Number: W81XWH-11-2-0225

TITLE: Integrative Lifecourse and Genetic Analysis of Military Working Dogs

PRINCIPAL INVESTIGATOR: William Kisseberth, DVM, MS, PhD

CONTRACTING ORGANIZATION: The Ohio State University
Columbus, OH 43210

REPORT DATE: December 2015

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
December 2015	Final	25 Sep 2011 - 24 Sep 2015

4. TITLE AND SUBTITLE

Integrative Lifecourse and Genetic Analysis of Military Working Dogs

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-11-2-0225

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
William Kisseberth, DVM, MS, PhD

E-Mail: kisseberth.2@osu.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Ohio State University
Columbus, OH 43210

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The purpose of this proposal is to provide insight into gene environment interactions. It leverages the simplified genetics and detailed records of the military working dog population. There are several critical aspects to meeting the aims of this proposal. 1) development of data driven selection criteria, 2) biological sampling of representative dogs, and 3) generation of mathematical methodologies capable of handling heterogenous data and statistical tests in consistent manner and providing clear and understandable results that are biologically valid. Here we provide a breakdown of the previous year's work and document our progress towards achieving the specific aims we proposed.

15. SUBJECT TERMS

Nothing listed

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	93	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

1. Introduction

The purpose of this proposal is to provide insight into gene environment interactions. It leverages the simplified genetics and detailed records of the military working dog population. There are several critical aspects to meeting the aims of this proposal. 1) development of data driven selection criteria, 2) biological sampling of representative dogs, and 3) generation of mathematical methodologies capable of handling heterogenous data and statistical tests in consistent manner and providing clear and understandable results that are biologically valid. Here we provide a breakdown of the previous year's work and document our progress towards achieving the specific aims we proposed.

2. A. Keywords

Canine, cancer, genetics, genomics, epigenetics, gene-gene interaction, gene-environment interaction, risk, longitudinal cohort

B. Abbreviations and acronyms

Military Working Dogs (MWDs), Principal Components Analysis (PCA), Multidimensional scaling (MDS), Joint Pathology Center (JPC), Armed Forces Institute of Pathology (AFIP), Animal Care and Use Review Office (ACURO), Institutional Animal Care and Use Committee (IACUC), Cooperative Research and Development Agreement (CRADA)

3. Accomplishments

A. What were the major goals of the project?

The proposal includes technical tasks to make the research possible, but these are not the major goals of the project. Because the present funded-project is first-of-its-kind in several respects, its research goals encompass more than the top single aim of understanding genetic and genetic/environmental risk factors for cancer in the MWDs. The main additional research goals are to establish the ideal analyses to be used and to understand their biostatistical properties and performance. The full goals have three elements:

i. **Creation of technological infrastructure**
ii. **Selection, development and validation of the most ideal analyses**
iii. **Execution of the final study of genetic and gene-environment risk factors for cancer in MWDs**

The first thus classifies as technical implementation whereas the second and third are research.

B. What was accomplished under these goals

i. Creation of technological infrastructure

- **Task 1: Regulatory approval:** regulatory approvals (IACUC, ACURO) and institutional agreements (CRADA) were received and executed, respectively. In this reporting period, an IACUC protocol set up exclusively to scan health care records (but involving no handling of animals) expired, and there has not been any scanning of records since then. This does not affect the separate current IACUC protocol for collection of biological samples. A new IACUC protocol to replace that for records scanning is pending.

- **Task 2: Data capture of veterinary records:** This task was directed by PI Alvarez, with support in the area of clinical data/health care records from PI Kisseberth.
  Task 2a) As planned, project manager and project programmer traveled to Lackland AFB, San Antonio, TX, to visit Military Working Dog (MWD) veterinary hospital in order to acquire information on data types, security, etc. and to establish protocols for data acquisition, scanning and

transfer between Lackland AFB and The Research Institute at Nationwide Children's Hospital (NCHRI), Columbus, OH. **2b)** As planned, regular communication has been maintained between NCHRI investigators and our veterinary technician at Lackland AFB. **2c)** Our Lackland AFB veterinary technician was originally planned to receive training on molecular methods so she could process the blood samples at Lackland AFB. This plan was deemed too risky to the quality of the biological samples and was changed to shipping fresh blood samples for processing at NCHRI. **2d)** We delivered to our veterinary technician at Lackland AFB and validated the computer and scanner to scan paper records into digital data for transfer to the DAPER database at NCHRI. NCHRI staff traveled to Lackland AFT to establish computational infrastructure as planned, but did not need to establish molecular infrastructure as originally planned (see 2c). **2e)** Ms. Perez and LTC Richard reviewed the available breeding program dogs available at Lackland AFB and meeting the study criteria. Drs. Kisseberth and Alvarez were also involved in this process. **2f)** Drs. Alvarez, Kisseberth and Zapata have evaluated possible exposures and biologically relevant events, and are prepared to focus such analysis more in depth in the final selected dogs. **2g)** Dr. Huang's programmer successfully created the DAPER database in which to collect the MWD health records. They completed the introduction of text-to-digital conversion software (ABBYY), and implemented custom forms for that conversion for one form and initiated that for pathology reports. In this past year, this digital-conversion work was put on hold to focus on more urgent needs for the project. It is anticipated that the necessary data capture can be done manually once the final set of dogs is chosen. To date, 257 veterinary medical records have been scanned, including for the dogs for which biological samples have been collected. An additional approximately 1,000 dogs have been identified that could serve for epidemiological analysis together with the 257 records already acquired. **2h)** Dr. Alvarez and Ms. Perez have continued evaluating progress and protocol compliance, in communication with the relevant Lackland AFB personnel. **2i)** Statistical analysis has continued, led by Drs. Zapata (mainly) and Alvarez. **2j)** Dr. Alvarez continued to supervise Ms. Perez in data handling, and initiated training of a new data handling assistant (Mr. I. Wolfe, Ohio State U. student; currently in unpaid volunteer status with academic credit for research).

- **Task 4 Identification, recruitment, and retention of cancer bearing and control dogs:** This task was directed by PI Alvarez, with support in the area of clinical data/health care records from PI Kisseberth. **Original Tasks 4a, b, c, d, e and f** all refer in part to non-inventory dogs at Lackland AFB. In our Year 2 Annual Report, we explained that MWDs from outside Lackland AFB were expected to not be necessary because acquisition of dogs with different environmental exposures was possible at MWD program at Lackland AFB and the extended MWD cohort housed at the Medina Annex, both in San Antonio.

  This task is well advanced, especially in the sense of acquisition of health care records from dogs prioritized by being cancer positive, high-cancer risk by family occurrence or of advanced age. The MWD subjects were selected from the full MWD population of Belgian Malinois. To date, 32 MWD blood samples were collected at Lackland AFB and shipped to Dr. Alvarez's lab (NCHRI, Columbus, OH). Collection is ongoing.

  Original Tasks 4a, b, d, e and f are technical, regarding setting up infrastructure and protocols. **Original Task 4c** refers to identification and ascertainment of dogs for the present study. A major advance in this area was the acquisition of the list of dogs for which there are pathology reports (i.e., processed by Joint Pathology Center (JPC)/Armed Forces Institute of Pathology (AFIP); those dogs are highly enriched for cancer-positive status. We were thus able to directly acquire the health records of those dogs, scan them and input to our database. Those records are for dogs that are not available

for collection of biological samples at Lackland AFB. They are useful for epidemiological information relevant to cancer in the MWD population, and to ultimately analyze together with that for the dogs used in the planned genetic studies.

- **Task 5: Molecular Characterization of cancer bearing and control dogs: Original Tasks 5a, b, c, d, and e** have two aspects – 1) molecular-technological development/implementation and validation, and 2) execution of the final molecular analysis once the full set of samples are collected. Notably, the latter cannot be done as the samples are collected in order to avoid experimental variability and batch effects.

  Detailed description and discussion of this task was provided in the first and second year annual reports. Those studies addressed implementation and validation of most molecular methodologies that we will use in this project. That included the critical molecular analyses that will be done first when all the biological samples have been collected. Most importantly, that includes the isolation and QC of high-quality DNA for SNP genotyping and DNA methylation analysis. We have also established SNP genotyping expertise and development of sequencing-based marker genotype validation and targeted DNA methylation assays that will be necessary following the GWA and DNA methylation analyses. The only development in this area in this reporting period was the collection of the majority of the biological samples acquired to date (current n=32).

- **Task 6 - Adaptation of existing resources, data storage and hosting:** This Task was directed by both PI's Huang and Alvarez, and the work was done by programmer J. Aaronson (who reported to PI Huang) working closely with PI Alvarez. **Tasks 6a, b, c and d** are technical – specifically listing each element of functionality to be added to the database of MWD health records and molecular data. In the two previous annual reports, we described the development of the virtual database server for this project and the implementation of software to convert paper health care records to digital. This reporting period we continued to import health care records data and to maintain the database.

- **Task 8**- Project management, Quality control and assurance, and Security: This task continues smoothly and without significant change as reported in the previous annual reports.

ii. Research goals: Selection, development and validation of the most ideal analyses (for execution of the final study of genetic and gene-environment risk factors for cancer in MWDs) – i.e. when the sample collection and molecular and statistical analyses are performed**)**

- **Task 3: "Development of Principal Components (PCs) and pedigrees"** Task 3 is advanced to the point it can be implemented once the final experimental MWD genotype data and familial relationship documentation are completed for the selected dogs. In this reporting period, we have begun collecting biological samples. These samples will be processed and will put through a pipeline of the methods we have developed in previous years.

  Original Tasks 3a, c, d: Principal Components Analysis (PCA) and pedigrees (3b was a related technical task, adding pedigree capabilities to MWD database): These all refer to methods and experimental design for understanding animal relatedness and population structure. This issue is unique to dog breeds and in particular to the MWDs because they are unlike other canine cohorts in their "evolutionary"/human-purposed selection/breeding (on top of the fact that dog breed population genetics are different than human and mouse, and are not uniform and generalizable because different breeds have vastly different population-genetic characteristics.

In the present reporting period we further developed our expertise in this area by analyzing the new dataset of Golden Retrievers with or without one of two cancer types (B cell lymphoma and hemangiosarcoma) from Tonomura et al. PMID: 25642983.This work is a good proxy for the present study because it i) represents multiple cancer types in one breed, and ii) it requires addressing population structure within an single breed. Included in these studies, we conducted GWA for each cancer separately, and both together (Appendix 1). We also formalized aspects of our understanding of canine breed genetics in the study design for an NIH R01 application that we submitted within the first quarter of this reporting period (this was included in the previous, year 3, annual report as Appendix 11).

Original Task 3e: Development for publication of technical and analytical novelties/ strengths of the current project for dissecting complex genetics and gene-environment interactions: In this reporting period, we continued to optimize and validate the biostatistical methods to determine genomewide significance of association. This is extremely important because it relates to the main challenges in complex genetics: mainly, multiple testing correction (and its associated loss of power to detect all associations but those of large effect sizes) and latent variables (e.g., population structure – addressed in previous section – and its resulting high false positive rate).

In this reporting period, we continued to formalize and optimize our methods developed in previous years – including the use of genomewide Intersection Union Testing association (or GIA). To do this, we expanded the types of populations analyzed – adding genotype datasets from the human HapMap resource (http://hapmap.ncbi.nlm.nih.gov/index.html.en) and from the largest dog genetics study to date (915 dogs from 80 breeds; Boyko et al., PMID: 20711490). We improved our ability to implement our analyses by developing new computer scripts that automate the analysis pipeline. We also developed and successfully implemented an analysis that is similar to versions of "identity by descent mapping" that have previously been used in human studies (manuscript #2 in bullet list below)

In the previous reporting period, we initiated epistasis analysis. That was continued in the present reporting period (Appendix 2, Figs. 1, 2). In addition, we developed new concepts and extensive statistical analyses regarding gene-environment interactions, and dog traits that could be associated with cancer risk (e.g., metabolic rate, lifespan and size). As we began discussing in the Q1 quarterly report and expanded on in later quarterly reports, new high impact publication suggests novel ways to think about the genetic and environmental contributions to cancer risk (Tomasetti C. and Vogelstein, B. Science 347:78-81; PMID: 25678653). In brief, they classified cancer as comprising (arbitrarily) two groups – Deterministic (significantly affected by genetic and environmental factors) and Replicative (dependent on stochastic somatic mutation resulting from the number of cell divisions, i.e., lifetime number of stem cells and their proliferation rate). The investigators note that in cancer epidemiology, "environmental" refers to anything not genetic. The stochastic processes associated with tissue development and maintenance are thus grouped with external environmental effects. This in turn suggests how those stochastic processes – in effect the total number of stem cells and their number of divisions – are likely to be affected by diverse genetic and environmental factors, beginning with those that influence size and weight. The information from this cited article can be used in this project to i) attempt the classification of military dog cancers into Deterministic or Replicative for selection of the best cancer types (most Deterministic), or ii) identify genetic or environmental contributions to Replicative cancer potential, such as mapped loci associated with dog size and weight (Appendix 3).

The development of this way of thinking of gene-environment interactions was a major area of productivity for our group in this reporting period (Appendices 3-6). Two manuscripts were

developed in this period (Appendices 5, 6) and submitted for publication a few months after the end of this reporting period.

- **Task 7: Pathway analysis and functional characterization.**
  Original Task 7a describes how the final analysis will use a blocking design to detect genetic and environmental effects once the final molecular genotyping is performed. **7b** describes how Intersection Union Testing/GIA will be used to analyze the final genetic and epigenetic molecular data generated. And **7c** and **d** describes how Kaplan Meyer analysis and Cox proportional hazard ratios will be calculated on the final data generated from molecular analysis of MWDs.

  Additionally, the overarching Task name and Tasks **7a**, **b** and **d** refer to pathways and other broad aspects such as bioinformatics and functional categories. As reported in the previous three annual reports, this task is advanced to the point that the planned analyses can be directly implemented on the final study data. In this reporting period, we continued to conduct broad cancer datamining studies that will be relevant to our analyses. These types of analyses are likely to dramatically enhance our canine studies, and could serve to support our findings. This is because there is vastly more human data available, especially in the area of somatic genetics and biochemical pathway/systems analysis.

  In this reporting period, Dr. Alvarez co-authored two relevant studies. In collaboration with Dr. G. Lorch and other OSU colleagues, Dr. Alvarez published a study of canine lung cancer (Clemente-Vicario et al. PMID: 26560147; Appendix 7 of this report). Dr. Alvarez's roles were to i) lead the Comparative Genomic Hybridization genetic assays and their analysis (which was done together with Dr. Jennie Rowell, OSU College of Nursing and former member of this project as a PhD student in Dr. Alvarez's lab); and ii) to conduct other analyses to determine the cancer relevance of observed alterations (mainly by datamining in human cancer genomics resources). In collaboration with Dr. Curtis Bird and his trainee, Dr. Alvarez also published an article on mammary cancer: Lutful Kabir, Alvarez and Bird, *Vet. Sci.* 2016, *3*(1), 1; doi:10.3390/vetsci3010001 (registering DOI; Appendix 8 of this report). In this work, Dr. Alvarez conducted the phylogenetic analysis and other bioinformatics and evolutionary analyses.

  In this reporting period, Dr. Alvarez was the communicating author on a comprehensive book chapter on canine translational genomics, as well as one of the authors of the introduction to the book (accepted for publication in 2016); Partnering PI W. Kisseberth and former project member Jennie Rowell are co-authors; note this is a high impact book series with wide circulation:

  Martic-Kehl, M., Festing, M. F. W., **Alvarez, C. E.**, Schubiger, P. A., Introduction to Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics; in Methods and Principles in Medicinal Chemistry series, Wiley-VCH, *In press*. (Appendix 9)

  Fenger J. M., Rowell, J. L., Zapata, I., London C. A., Kisseberth W. C.**, Alvarez, C. E.** (2015) Dog models of naturally occurring cancer; in Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics Methods and Principles; Medicinal Chemistry series, Wiley-VCH, *In press*. (Appendix 10)

  This work includes a section titled "Necessary developments for realizing the potential of canine models", with subsection "Epidemiology, longitudinal cohorts, tissue repositories and integrative genomics". This section brings attention to the approaches behind the present effort and cite this study through the DoD CDMRP annual report. That section concludes with the following: "An ideal longitudinal cohort should make it possible to gather the following data in electronic form: clinical, environmental, pedigree, genetic and molecular phenotypes."

C. What opportunities for training and professional development has the project provided?

Dr. Isain Zapata has trained as a postdoctoral fellow through this reporting period. He conducted the statistical genetic analyses reported here for Task 3. He was first author on two manuscripts developed in this period and submitted shortly after (Appendices 5, 6). He also presented posters on that work at two internal events, as reported in Q3 and Q4 quarterly report appendices (Appendix 4).

Dr. Alvarez continues to formally mentor Dr. Jennie Rowell (former project member as PhD student in Dr. Alvarez's lab). She is now a tenure track faculty member in the OSU College of Nursing. In this reporting period, Dr. Rowell was second author on the book chapter described in B.ii.Task 7, last paragraph. Dr. Alvarez also was a member of a working group for a grant application being developed by Dr. Rowell with colleagues in her college.

D. How were the results disseminated to communities of interest?

We have begun to draw attention to this project in publications and international presentations (Symposium: Cancer Drug Development – Predictability of Animal Experiments, Collegium Helveticum, Zurich, Switzerland). That in turn resulted in an invited book chapter below, which is described in B.ii.Task 7, last paragraph.

i. We will draw attention to this project in the following book-series publication that include former project member Dr. J. Rowell; Partnering PI Dr. W. Kisseberth; and as communicating author, Lead PI Dr. Alvarez:

Martic-Kehl, M., Festing, M. F. W., **Alvarez, C. E.**, Schubiger, P. A. (2016) Introduction to Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics; in Methods and Principles in Medicinal Chemistry series, Wiley-VCH, *In press*. (Appendix 9)

Fenger J. M., Rowell, J. L., Zapata, I., London C. A., Kisseberth W. C.**, Alvarez, C. E.** (2016) Dog models of naturally occurring cancer; in Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics; in Methods and Principles in Medicinal Chemistry series, Wiley-VCH, *In press*. (Appendix 10)

ii. We have also drawn attention to this project at local university presentations:

OSU Humanities & Cognitive Sciences High School Summer Institute, lecture on behavioral and canine genetics (Aug 13, 2015) (Appendix 11)

OSU VME 6540 Structure & Function of Cells (Cell Biology), Lecture on canine genetics & genomics (Oct 5, 2015) (Appendix 12)

iii. The following articles on canine cancer were published:

Lutful Kabir, F.M., Alvarez, C.E. and Bird, R.C., Canine Mammary Carcinomas: A Comparative Analysis of Altered Gene Expression, *Vet. Sci.* 2016, *3*(1), 1; doi:10.3390/vetsci3010001 (registering DOI) (Appendix 8)

Clemente-Vicario F, Alvarez CE, Rowell JL, Roy S, London CA, Kisseberth WC, Lorch G. Human Genetic Relevance and Potent Antitumor Activity of Heat Shock Protein 90 Inhibition in Canine Lung Adenocarcinoma Cell Lines. PLoS One. 2015 Nov 11;10(11):e0142007. doi: 10.1371/journal.pone.0142007. eCollection 2015. PubMed PMID: 26560147 (Appendix 7)

E. What do you plan to do during the next reporting period to accomplish the goals?

i. We plan to publish some of the work previously submitted and under revision (GIA methods manuscript), and we plan to submit other articles, including those mentioned as manuscript #1 and #2 above (Task B.ii.Task 3).

ii. We plan to complete the collection of MWD biological samples and conduct the planned molecular genetic studies.

4. **Impact**

A. What was the impact on the development of the principal discipline(s) of the project?

Our publication Alvarez, CE, ILAR J 2014 has already been prominently referenced:

- This citation is important because it is in a high impact journal read by the broad field of biomedical research: Science Translational Medicine: Kol A, Arzi B, Athanasiou KA, Farmer DL, Nolta JA, Rebhun RB, Chen X, Griffiths LG, Verstraete FJ, Murphy CJ, Borjesson DL. Companion animals: Translational scientist's new best friends. Sci Transl Med. 2015 Oct 7;7(308):308ps21. doi: 10.1126/scitranslmed.aaa9116. Review. PubMed PMID: 26446953.
- In "Comparative oncology: what dogs and other species can teach us about humans with cancer". Schiffman JD, Breen M. Philos Trans R Soc Lond B Biol Sci. 2015 Jul 19;370(1673). pii: 20140231. This publication cites us (ref. #3) in the Introduction, which starts "Comparative oncology is a quickly expanding field … The study of naturally occurring cancers in the domestic dog provides a suitable model for advancement of the understanding, diagnosis and management of cancer in humans [1–4]."
- In "The Golden Retriever Lifetime Study: establishing an observational cohort study with translational relevance for human health". Guy MK, Page RL, Jensen WA, Olson PN, Haworth JD, Searfoss EE, Brown DE. Philos Trans R Soc Lond B Biol Sci. 2015 Jul 19;370(1673). pii: 20140230. These investigators wrote about our paper (ref#9): "Remarkable progress has been made in the last 10 years, … to begin to investigate the multifactorial, complex nature of … cancer development in dogs. A recent review provides an interesting and salient assessment of this area [9]."

We also expect that our newly developed statistical methods (mainly GIA) and trait mapping studies (submitted manuscripts #1 and #2 mentioned above) will have a strong impact on the field of genetics.

B. What was the impact on other disciplines?

Our work has drawn attention from the fields of comparative biology and animal models of human disease. While the latter is clear in the area of cancer genetic mapping, there are also important aspects regarding epidemiology, human disease relevance and pharmacology. The appreciation of these is demonstrated by the PI being invited to contribute a chapter to a book largely dedicated to reproducibility in research using animal models to develop and validate new pharmacological therapies, in the Wiley Methods and Principles in Medicinal Chemistry series (Fenger et al. 2016 above; Appendix 10).

C. What was the impact on technology transfer?

Nothing to report

D. What was the impact on society beyond science and technology?

There is a high level of interest in MWDs in the US population at large, especially in military families that understand the MWDs are the most effective way to mitigate the major cause of morbidity and mortality in deployed military personnel – exposure to explosions. We predict the present project to understand cancer in that population will gain media and popular attention.

5. **Changes/problems**

Nothing to report

6. **Products**

A. Publications, conference papers, and presentations

Martic-Kehl, M., Festing, M. F. W., **Alvarez, C. E.**, Schubiger, P. A. (2016) Introduction to Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics; in Methods and Principles in Medicinal Chemistry series, Wiley-VCH, *In press*. (Appendix 9)

Fenger J. M., Rowell, J. L., Zapata, I., London C. A., Kisseberth W. C.**, Alvarez, C. E.** (2016) Dog models of naturally occurring cancer; in Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics; in Methods and Principles in Medicinal Chemistry series, Wiley-VCH, *In press*. (Appendix 10)

Lutful Kabir, F.M., Alvarez, C.E. and Bird, R.C., Canine Mammary Carcinomas: A Comparative Analysis of Altered Gene Expression, *Vet. Sci.* 2016, *3*(1), 1; doi:10.3390/vetsci3010001 (registering DOI) (Appendix 8)

Clemente-Vicario F, Alvarez CE, Rowell JL, Roy S, London CA, Kisseberth WC, Lorch G. Human Genetic Relevance and Potent Antitumor Activity of Heat Shock Protein 90 Inhibition in Canine Lung Adenocarcinoma Cell Lines. PLoS One. 2015 Nov 11;10(11):e0142007. doi: 10.1371/journal.pone.0142007. eCollection 2015. PubMed PMID: 26560147 (Appendix 7)

Zapata, I., Alvarez, C.E. Genome Mapping of Aggression and Fear Traits in Dogs. The Research Institute at Nationwide Children's Hospital Poster Presentation, April 2015

Zapata I., Serpell, J.A., Alvarez, C.E., Genetic mapping of canine fear and aggression. The Research Institute at Nationwide Children's Hospital, Annual Retreat (poster presentation) November 2015 (Appendix 4)

OSU Humanities & Cognitive Sciences High School Summer Institute, lecture on behavioral and canine genetics (Aug 13, 2015) (Appendix 11)

OSU VME 6540 Structure & Function of Cells (Cell Biology), Lecture on canine genetics & genomics (Oct 5, 2015) (Appendix 12)

B. Website(s) or other Internet site(s)

Nothing to report

C. Technologies or techniques

Further developed GIA statistical method for genomewide association of genetic or genomic data; manuscript was submitted and is under revision. Developed and implemented an adaptation of identity by descent mapping; manuscript was submitted shortly after end of this reporting period (referred to as manuscript #2 above).

D. Inventions, patent applications, and/or licenses

Nothing to report

E. Other Products

We developed and validated a new computer script to automate an analysis pipeline that constitutes our GIA methodology

7. Participants & other collaborating organizations

A. What individuals have worked on the project?

Carlos E. Alvarez, PhD, unchanged
Kun Huang, PhD, unchanged
Bill Kisseberth, DVM, PhD, unchanged
Jacob Aaronson (programmer reporting to PI Huang), unchanged
Isain Zapata, PhD (postdoctoral fellow), unchanged
Michelle Garcia (formerly Perez), RVT, unchanged

B. Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to Report

C. What other organizations were involved as partners?

Dr. J. Serpell of University of Pennsylvania is now a collaborator on our new studies that were submitted for publication and referred to above as manuscript #1 (Appendix 5).

8. Special reporting requirements

This is a Collaborative Award and the report is written to document all three investigators' roles. All work was done by PI Alvarez and his staff except where noted in the header that another PI and their staff were involved (Huang, Task 6; and Kisseberth, Tasks 2 and 4).

9. Appendices:

1) New Intersection Union Test genomewide association analysis implementations: i) multi-cancer-type GWA in one breed & ii) single breed population structure
2) Continuation of refinement of genomewide analysis: i) epistasis & ii) relationship between size and lifespan
3) Consideration of new ways of thinking about gene-environment interactions
4) Analysis of novel gene-environment interactions possibilities: Poster presentations: Zapata et al. Genetic mapping of canine fear and aggression
5) Analysis of novel gene-environment interactions possibilities: Draft manuscript: Zapata, Serpell and Alvarez: Genetic mapping of canine fear and aggression
6) Analysis of novel gene-environment interactions possibilities: Zapata and Alvarez: Fine-mapping and biological relevance implicate *GNAT3* and *IGSF1* in canine fear and aggression
7) Published article: Human Genetic Relevance and Potent Antitumor Activity of Heat Shock Protein 90 Inhibition in Canine Lung Adenocarcinoma Cell Lines
8) Published article: Canine Mammary Carcinomas: A Comparative Analysis of Altered Gene Expression
9) Accepted manuscript draft: Introduction to Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics
10) Accepted manuscript draft: Dog models of naturally occurring cancer; in Animal Models for Human Cancer
11) Presentation: OSU Humanities & Cognitive Sciences High School Summer Institute
12) Academic lecture/curriculum development: OSU VME 6540 Structure & Function of Cells (Cell Biology), Lecture on canine genetics & genomics

Appendix 1

New Intersection Union Test genomewide association analysis implementations: i) multi-cancer-type GWA in one breed & ii) single breed population structure

We initiated reanalysis of a new dataset that represents an attractive proxy of the military dog data we will ultimately analyze:

> Tonomura N, Elvers I, Thomas R, Megquier K, Turner-Maier J, Howald C, Sarver AL, Swofford R, Frantz AM, Ito D, Mauceli E, Arendt M, Noh HJ, Koltookian M, Biagi T, Fryc S, Williams C, Avery AC, Kim JH, Barber L, Burgess K, Lander ES, Karlsson EK, Azuma C, Modiano JF, Breen M, Lindblad-Toh K. Genome-wide Association Study Identifies Shared Risk Loci Common to Two Malignancies in Golden Retrievers. PLoS Genet. 2015 Feb 2;11(2):e1004922. doi: 10.1371/journal.pgen.1004922. eCollection 2015 Feb. PubMed PMID: 25642983

This data is ideal because, unlike previous available data, it examines more than one type of cancer (B-cell lymphoma and hemangiosarcoma) in a single cohort or breed (Golden Retriever; approximately 180 affected and 170 controls). This dataset is thus analogous to our Belgian Malinois pan-cancer study. Because our new method (GIA, genomewide intersection union test association) is well suited to single marker analysis, our initial studies have been to compare such analysis to the conventional analysis reported in that study ("collapsing method" whereby multiple testing burden is dramatically reduced by using linkage disequilibrium-defined haplotype blocks instead of single SNP markers). A second aspect of our reanalysis addresses the distinguishing characteristic of GIA – the fact that its statistical mechanism is akin to conducting meta-analylsis (of subsets of a single dataset). Our first analysis shown in Appendix 1 Fig. 1 begins to address both of those aspects using conventional GWAS analysis in PLINK – i.e., 1) single marker analysis of the two cancer types separately, and 2) combined analysis whereby both cancer types are combined in a cancer vs. normal GWAS.

Appendix 1. Figs. 1, 2, 3. Mapping 2 cancers in one breed

Golden Retrievers

BLSA, B cell lymphoma

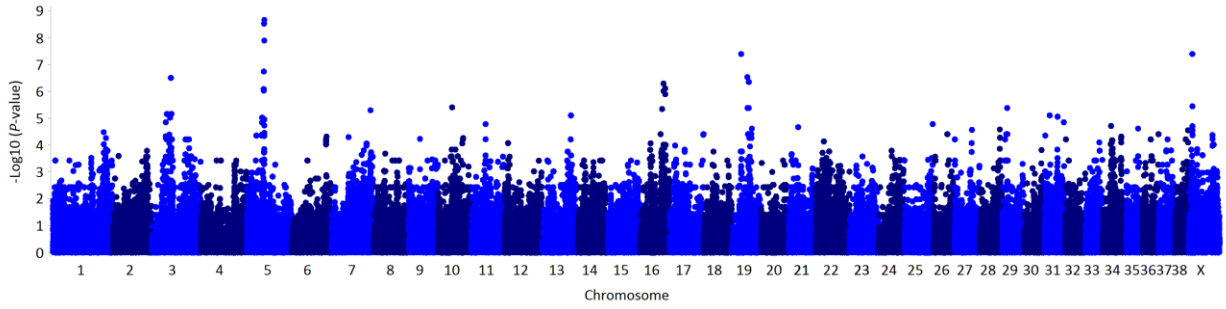
HAS, hemangiosarcoma

Arrows denote upstream or downstream of a peak SNP

Analysis of genomewide association using conventional methods (in PLINK)

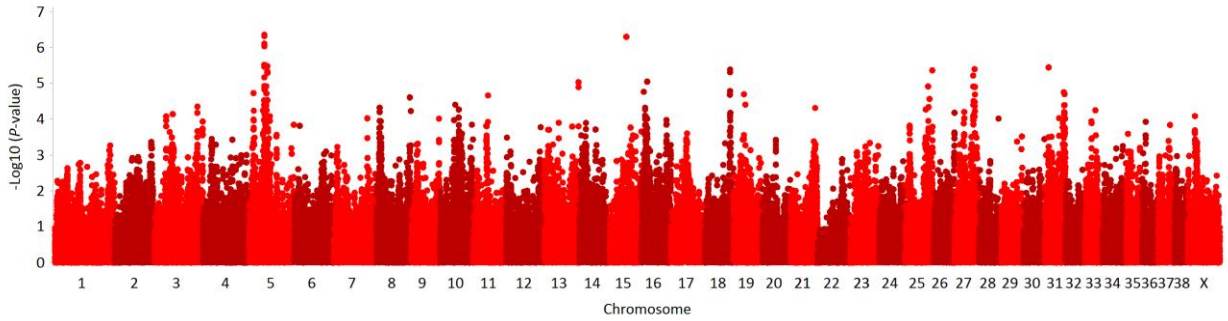
Dataset from Tonomura et al. PMID: 25642983

BLSA



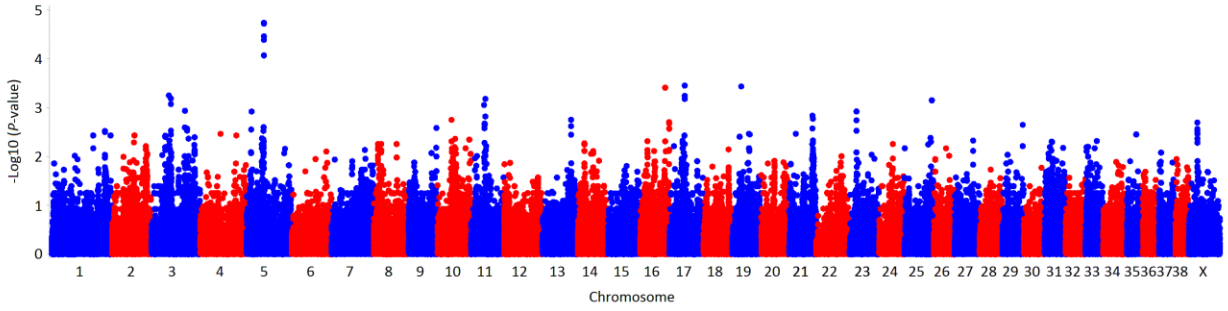
Chr	bp	P	Gene
5	36839546	2.14E-09	ARHGAP44
X	5487485	3.95E-08	KAL1
19	23139076	3.97E-08	↓IWS1, ↑PROC
16	51008854	5.02E-07	↓UBE2F, ↑DNMT1
10	36123523	3.88E-06	↓FARSB?, ↑PPKRA

HAS



Chr	bp	P	Gene
5	32901346	4.32E-07	ALOX12B
15	39512593	4.93E-07	↓SLC17A8, ↑NR1H4
5	38808874	3.24E-06	↓UBL5, ↑PMP22
31	10160234	3.52E-06	↓ROBO2, ↑ENOPH1
27	40209830	3.95E-06	NDUFA9

BLSA-HAS



Chr	bp	P	Gene
5	36848237	1.81E-05	ARHGAP44
17	32453830	0.000351	↓SLC8A1
19	23139076	0.000366	↓IWS1, ↑PROC
16	54230145	0.000387	↓USP17L1, ↑ARHGEF10
3	41219385	0.00056	↓MEF2A, ↑LRRC28

Appendix 2

Continuation of refinement of genomewide analysis: i) epistasis & ii) relationship between size and lifespan

In this reporting period, we continued refining and validating our novel methods to conduct genetic (genomewide association analysis or GWAS) and multidimensional analysis: Genomewide Intersection-Union-Testing (IUT) Analysis (GIA). The latest efforts in this area are described with figures in Appendix 1. In this reporting period, we initiated studies of new datasets that provide new opportunity to test and validate our methods. Those include human genetic data from HapMap (http://hapmap.ncbi.nlm.nih.gov/index.html.en). This is desirable because the extent of linkage disequilibrium (LD) is reduced in the range of 50-fold compared to humans; that allows us to test our approach on the other end of the spectrum from pure breed data with exaggerated LD. At the same time, that data has the benefit of being dramatically stratified like dog breeds, because it is made up of genotypes from select human populations throughout the world – some closely related, others distantly related. Unlike most human genetic data, this dataset is not highly regulated because of the anonymous design of the project. In addition, we requested and received the dataset from the largest dog genetic study published to date (PLoS Biol. 2010 Aug 10;8(8):e1000451.): 915 dogs from 80 breeds. This allows for a different set of partially overlapping dogs as reported by Vaysse et al. (LUPA consortium, cited above), and including several of the same traits. The ongoing analyses are encouraging (see Appendix 2, Figures and text).

In addition to developing and optimizing the statistical methods (Appendices 1, 2), we have initiated the design of improvements of the computer script to automate the GIA analysis. The script has been validated and implemented by repeating previous analyses. Going forward, we plan to optimize it further and add functionalities that will further accelerate our GIA analyses – including creation of an interface that allows selection of variable parameters. For example the current version of the automated script can only analyze case-control data; we thus plan to add the option of analyzing quantitative trait loci (QTL).

In this reporting period, we initiated discussions of analysis options for gene-environment interactions. In addition, we recognize that a similar effect can pertain to non-cancer clinical traits or disease. This can be important because there can be both positive or negative correlations between cancer and such traits/diseases. And some of those relationships can be predicted to have effects on cancer rate. Examples include age- or metabolism- related traits that indicate advanced aging, or obesity or larger body size of some dogs vs. others (Appendix 2, Figs. 3, 4).

Epistasis analysis

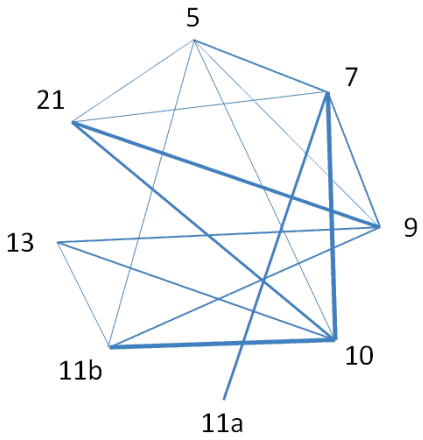
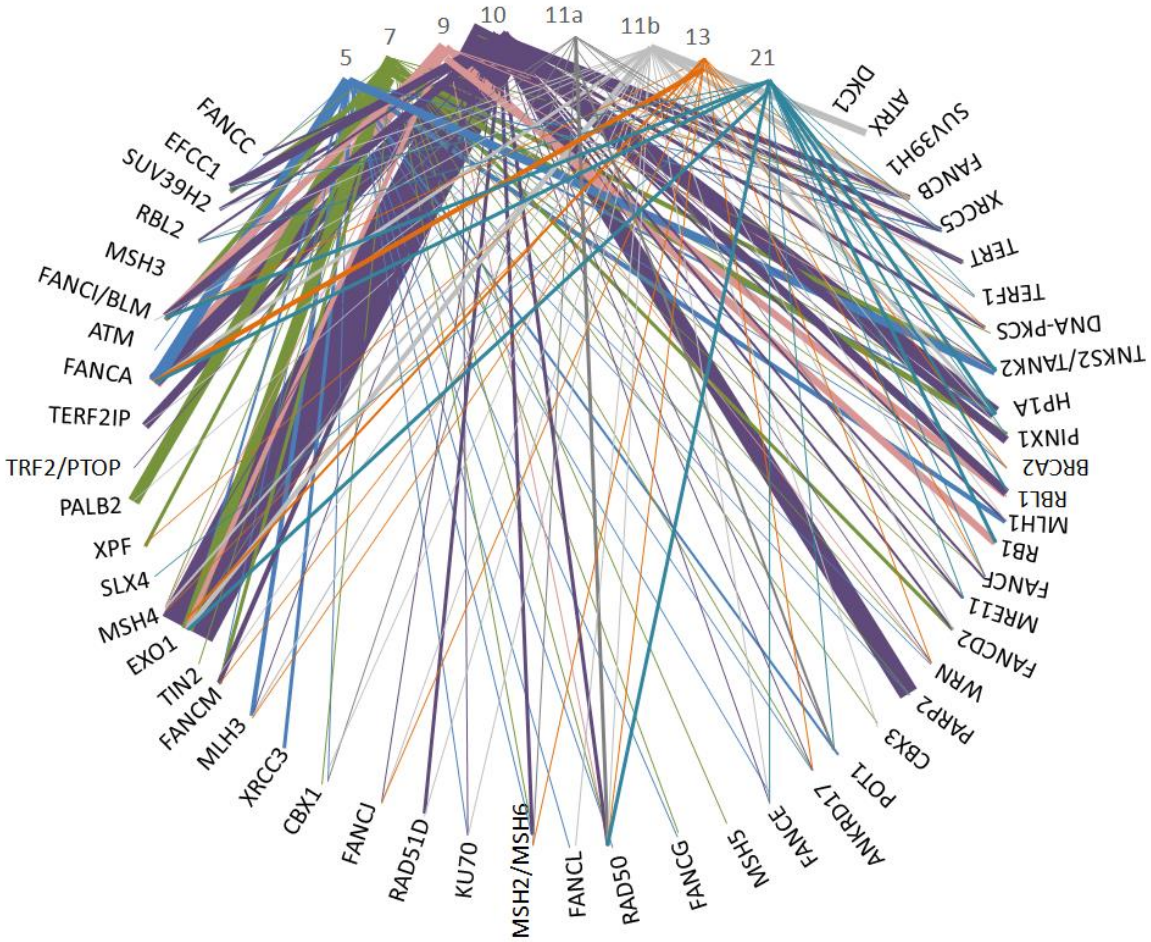


Figure 1 (right). Epistatic interactions between individual SNPs within the relevant lifespan-associated QTLs identified through intersect union test GWAS. This plot shows the number significantly correlated SNPs within the relevant locus. The more interactions there are (the thicker the line) the more correlation exists among the locus. Higher correlation may be suggestive of a interaction of effects. Within hits, the one in chromosome 10 appears to be the largest one. Data source: Vaysse et al. 2011.

Figure 2 (bottom). Epistatic interactions between the individual SNPs within the relevant lifespan associated QTLs identified though intersect union test GWAS against a list of genes involved in telomere regulation. Telomere length has been shown to be highly predictive of life expectancy. This plot evaluated the epistatic interactions between the lifespan associated QTLs and telomere regulation genes. The thicker the line, the larger the correlation that exists. The most relevant epistatic association detected was the hit in chromosome 10, which has several associations that are evident but especially to EXO1.



Interbreed Size vs Lifespan association.

Two main studies have been published in the past where the interbreed size variation was evaluated in a GWAS setting. Those studies are Boyko et al. 2010 and Vaysse et al. 2011.Both studies have similar conclusions in regards to size but none of them has been evaluated for lifespan QTL associations. We finally gained access to both datasets so they can be compared. Both datasets were evaluated originally using different methodologies that are not easy to compare. We began by analyzing the traits of Size and Lifespan using our intersect union test GWAS. All QTLs are evaluated on stereotypical phenotypes o each breed.

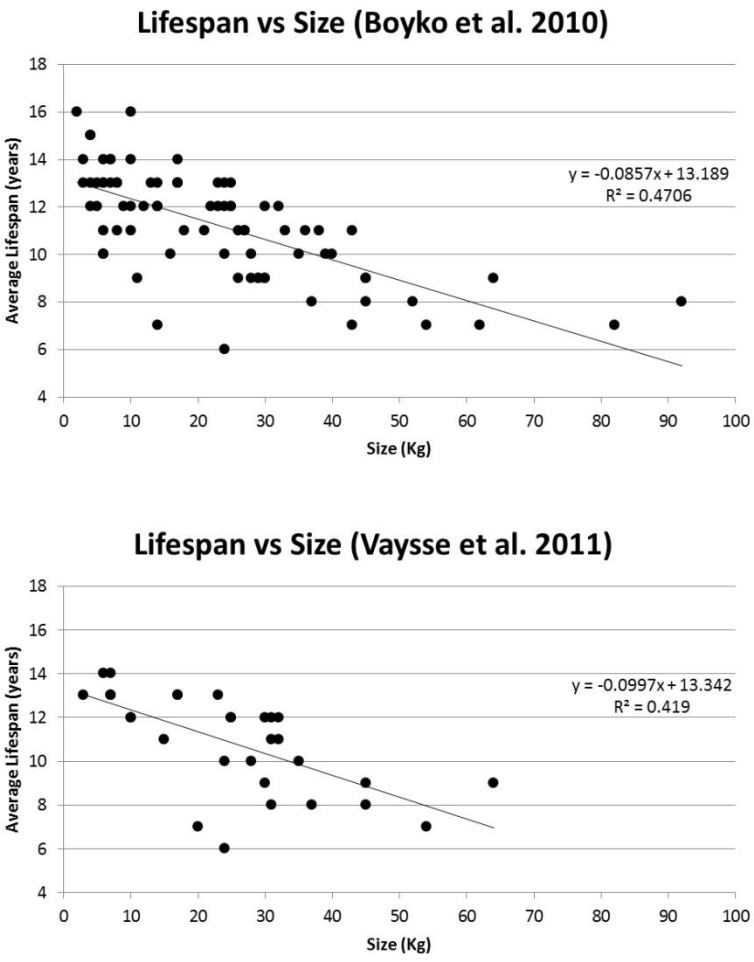


Figure 3 (right). Empirical linear regression curves were constructed to evaluate the distribution of the selected breeds to be compared. Each study used a different set of breeds. Boyko et al. 2010 included 80 breeds represented by approximately 10 individuals per breed. While Vaysse et al. 2011 included 30 breeds represented by approximately 16 individuals per breed. The datasets only contained 19 breeds that were common in both studies. The empirical linear regression equation and R-squared value are very close to each other even when the breed population is different. This evaluation suggests the presence of some outliers that may be negatively impacting the QTL detection efficiency. (i.e. the English Bulldog has a very short lifespan for what was expected for his weight). It will be further evaluated if the removal of outlier breeds improves the efficiency of the detection.

Figure 4 (Bottom). Manhattan plots of Size and Lifespan QTLs using the 4 group intersect union test GWAS. In the case of Size, the main hit in chromosome 15 and the X chromosome become quite evident in both datasets; however, the Vaysse dataset detected some addional signal coming from chromosomes 4, 7 and 10 that are not that evident the Boyko dataset. Boyko dataset included several lager breeds that are not represented in the Vayse dataset. That feature may be hiding some of the other relevant hits. In terms of Lifespan, hits are not that evidently clear as compared to size, there is no main hit that is common in both datasets besides chromosome 5 and 10. In this case it may also be important to evaluate if the breed population representation can be altering the findings.

Appendix 3

Consideration of new ways of thinking about gene-environment interactions

In this reporting period, we have devoted significant effort to considering a new way of thinking about gene-gene and gene-environment questions in our present study. A recent high-impact research article proposed a new way of thinking about cancer risk; this has generated a great deal of discussion:

Tomasetti C, Vogelstein B. Cancer risk: role of environment—response. Science. 2015 Feb 13;347(6223):729-31. doi: 10.1126/science.aaa6592. Epub 2015 Feb 5. PubMed PMID: 25678653.

> "*If hereditary and environmental factors cannot fully explain the differences in organ-specific cancer risk, how else can these differences be explained? Here, we consider a third factor: the stochastic effects associated with the lifetime number of stem cell divisions within each tissue. In cancer epidemiology, the term "environmental" is generally used to denote anything not hereditary, and the stochastic processes involved in the development and homeostasis of tissues are grouped with external environmental influences in an uninformative way. We show here that the stochastic effects of DNA replication can be numerically estimated and distinguished from external environmental factors. Moreover, we show that these stochastic influences are in fact the major contributors to cancer overall, often more important than either hereditary or external environmental factors.*
>
> *That cancer is largely the result of acquired genetic and epigenetic changes is based on the somatic mutation theory of cancer (9–13) and has been solidified by genome-wide analyses (14–16). The idea that the number of cells in a tissue and their cumulative number of divisions may be related to cancer risk, making them more vulnerable to carcinogenic factors, has been proposed but is controversial (17–19)…"*

Given the focus of our study on identifying genetic and environmental contributions to cancer risk in the military dog population, that way of thinking about cancer risk presents some strong candidates for gene-environment interactions that could be relevant. Those include dog size (or growth rate, duration of growth period, or other metabolic traits), aging rate (as cancer is associated with advanced age) and aggression (because the military dogs are bred for aggression, aggression is associated with stress which is associated with cancer, and it represents a possible gene/amimal-environment (incl humans and other dogs) interaction). In this reporting period, we conducted analyses on cross-breed association of aggression and genetic markers (using

published datasets of breed genotypes (Boyko et al. PMID: 20711490; Vaysse et al. PMID: 22022279) and phenotypes (c-BARQ resource; McGreevy et al. PMID: 24358107)). Strikingly, we were able to map genomic loci associated with aggression across diverse dog breeds (Appendix I). Because the analysis is across breeds, the mapping is not subject to the extreme linkage disequilibrium distances within dog breeds – and thus serves as fine mapping to strongly implicate single genes.

To refine the aggression trait itself, we first conducted Principal Components Analysis (PCA) of broad behavioral traits (as breed stereotypes). This revealed a cluster of three related traits – dog-directed aggression, human stranger-directed aggression and fear of human stranger. As a result, we predicted that genetic mapping of those traits separately could yield loci shared by those traits. Strikingly, that is what we found (Appendices 4-6). Moreover, the genes implicated – *GNAT3* and *IGSF1* – are exclusively dedicated to sensory perception (we hypothesize in pheromone sensation in the vomeronasal organ or olfactory epithelium) and the hypothalamic-pituitary axis that drives the stress response. We were able to show this effect in two genotype datasets with different breeds, and to establish that it is associated with the three aggression/fear traits mentioned above (but not universally across other behavioral traits). These findings are very encouraging because they suggest 1) aggression can be mapped, and 2) this particular cluster of aggression/fear traits are explicitly associated with gene-environment interactions (environment being non-familiar humans/animals). Thus these traits appear to be highly relevant to our study because the dogs are bred and trained for aggression.

Appendix 4

Analysis of novel gene-environment interactions possibilities
Poster presentations: Zapata et al.

Explained in Appendix 3 text.

Poster presentations of this work:

Zapata, I., Alvarez, C.E. Genome Mapping of Aggression and Fear Traits in Dogs. The Research Institute at Nationwide Children's Hospital, April 2015

Zapata I., Serpell, J.A., Alvarez, C.E., Genetic mapping of canine fear and aggression. The Research Institute at Nationwide Children's Hospital, Annual Retreat, November 2015 (poster media shown on next page)

Genome Mapping of Aggression and Fear Traits in Dogs

Isain Zapata & Carlos E. Alvarez

Center for Molecular and Human Genetics, The Research Institute at Nationwide Children's Hospital, Columbus, OH

Introduction

Genomewide genetic associations (GWAS) of behavioral traits is difficult in humans. This is due to heterogeneity, biological complexity, and the challenge of phenotyping large numbers of individuals. Thus, little is known about the genetics and molecular biology of behavior.

The genetic history of dogs has resulted in hundreds of isolated populations with diverse morphological and behavioral traits. Those various breed stereotypes can be relatively easily ascertained through i) direct phenotyping of many individuals from diverse breeds, or ii) questionnaire-based surveys of very large numbers of diverse pedigree dogs.

We are identifying common gene variations that are associated with different behaviors (such as aggression and fear) across dog breeds. We propose that knowledge of their genetic basis will enable us to dissect many of these behaviors and their correlations. That new understanding will lead to novel strategies for improving human mental health.

Objective

Identify relevant genome regions associated with behavioral traits of aggression and fear in the dog.

Methods

Phenotypes

- C-BARQ values. Behavioral means by breed of the 30 most popular breeds by AKC. (Serpell *et al*. 2014)
- Owner reported behavioral data of AKC registered dogs.
- ~175 subjects per breed. (Min: 49-Bulldog Max:1120-Labrador Retriever)
- Aggression decomposed into 4 classes
- Fear decomposed into 5 classes

Datasets

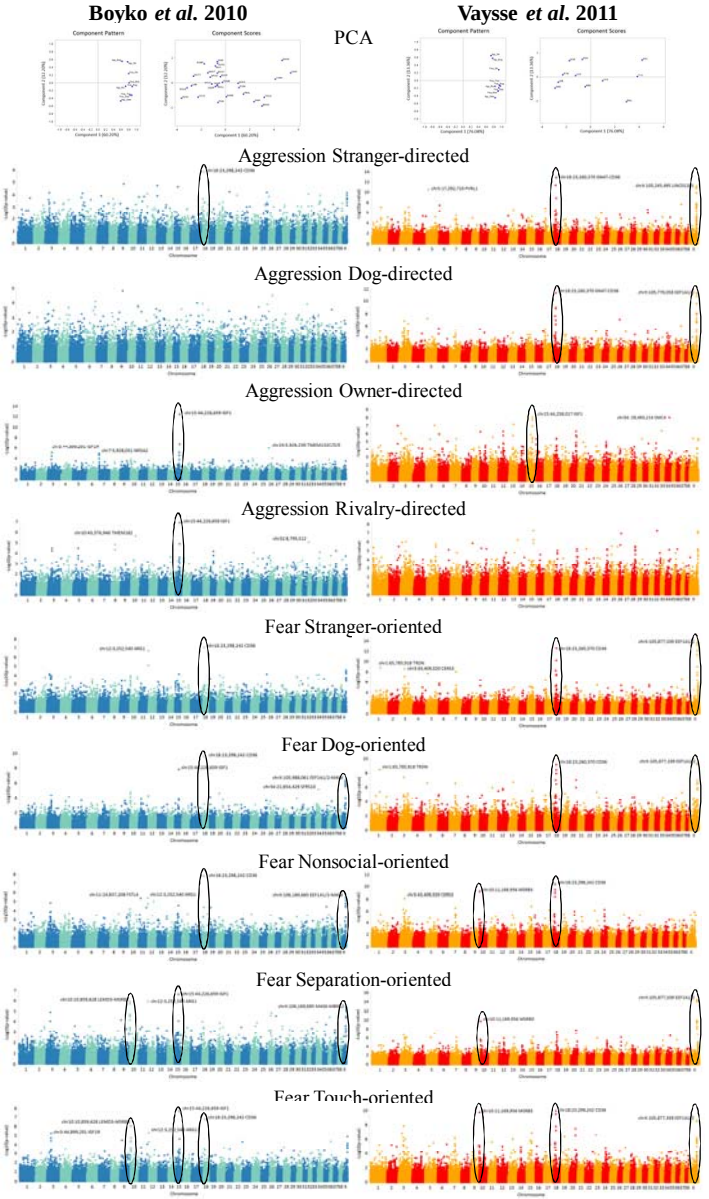
- Boyko dataset: 327 subjects from 29 Breeds, ~49,000 SNP Affymetrix v.2 Canine array (Boyko *et al*. 2010).
- Vaysse dataset: 150 subjects from 11 breeds, ~178,000 SNP Illumina Canine HD array (Vaysse *et al*. 2011).

Analysis

- All descriptive statistics performed on SAS v.9.3.
- GWAS performed using GEMMA v.0.94.1 (Zhou *et al*. 2012).
- Population structure removed by centered relatedness matrix correction.
- Association test performed using Univariate Linear Mixed Model using the likelihood ratio test.

➢ Genomewide significance P-value cutoff at <1E-5 for the Boyko dataset & <1E-8 for the Vaysse dataset.

Results



Findings

Region	Trait	Genes in region§	Remarks
chr10:11.1Mb	Fear-Nonsocial, Fear-Separation, Fear-Touch	MSRB3, ↓LEMD3, ↑HMGA2	Domestication associated region
chr15:44.2Mb	Aggression-Owner, Agression-Rivalry, Fear-Separation, Fear-Touch	IGF1	Well known small dog size allele
chr18:23.2Mb	Aggression-Stranger, Agression-Dog, Fear-Stranger, Fear-Dog, Fear-Nonsocial, Fear-Touch	CD36, ↑GNAT3	Si and Di region
chrX:105.8Mb	Aggression-Stranger, Aggression-Dog, Fear-Stranger, Fear-Dog, Fear-Nonsocial, Fear-Separation, Fear-Touch	IGSF1, ↑FIRRE, ↑MASK	Si and Di region

- § Genes are not necessarily biologically causative.
- Traits in bold were detected in both datasets.

Conclusions

- ***Aggression and fear-predisposition are correlated – both by PCA and GWAS***
- ***We identified four distinct regions highly associated with aggression and fear, and have a solid gene candidate for three of those***

References

- Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, Lohmueller KE, Zhao K, Brisbin A, Parker HG, vonHoldt BM, Cargill M, Auton A, Reynolds A, Elkahloun AG, Castelhano M, Mosher DS, Sutter NB, Johnson GS, Novembre J, Hubisz MJ, Siepel A, Wayne RK, Bustamante CD, & Ostrander EA. 2010. A simple genetic architecture underlies morphologic variation in dogs. *PLoS Biol*. 8(8):e1000451.
- Serpell, J. & D. Duffy. 2014. Dog Breeds and Their Behavior. Domestic Dog Cognition and Behavior. A. Horowitz, Editor. Springer-Verlag Berlin Heidelberg: 31-57.
- Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Rosengren Pielberg G, Sigurdsson S, Fall T, Seppälä E, Hansen M, Lawley C, Karlsson E, Bannasch D, Vilà C, Lohi H, Galibert F, Fredholm M, Häggström J, Hedhammar On, André C, Lindblad-Toh K, Hitte C, & Webster M. 2011. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genetics*. 7(10):3
- Zhou X, & M Stephens. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*. 44: 821-824.

Appendix 5

Analysis of novel gene-environment interactions possibilities: Draft manuscript: Zapata, Serpell and Alvarez

See discussion of rationale for relevance to cancer risk and this project in Appendix 3.

Title: Genetic mapping of canine fear and aggression

Authors: Isain Zapata[1], James A. Serpell[2] and Carlos E. Alvarez[1,3,4]*

Affiliations:
[1]Center for Molecular and Human Genetics, The Research Institute at Nationwide Children's Hospital, Columbus, OH, 43205, USA
[2]Center for the Interaction of Animals and Society, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA
[3]Department of Pediatrics, The Ohio State University College of Medicine, Columbus, OH, 43210, USA
[4]Department of Veterinary Clinical Sciences, The Ohio State University College of Veterinary Medicine, Columbus, OH, 43210, USA
*Corresponding Author: alvarez.73@osu.edu

One Sentence Summary: Genome scanning across dog breeds reveals that different genome loci are associated with two clusters of fear and aggression traits.

Abstract:
Little is known about the genetics of fear and aggression in humans and other mammals. We conducted genomewide association (GWA) mapping of breed stereotypes for many fear and aggression traits across several hundred dogs from diverse breeds. Those findings were confirmed using a second cohort of partially overlapping breeds. Lastly, we used the validated loci to create a model that effectively predicted fear and aggression stereotypes in a third group of dogs from breeds that were not involved in the mapping studies. Our major findings are that i) owner directed aggression and dog rivalry are associated with variation for small body size, and ii) two loci, on chromosomes 18 and X, are associated with several traits, including fear and aggression that is directed toward dogs and human strangers. We discuss the biological relevance of genes at the GWA peaks and the possibility that those variants were involved in the domestication process.

Main Text:

Introduction

It is difficult to perform genomewide genetic association studies (GWAS) of human behavior. This is due to heterogeneity, biological complexity, ambiguous phenotype classifications and the challenges of phenotyping large numbers of individuals. As a result, very little is known about human behavioral genetics. Most of the progress has been driven by the availability of epidemiological data of medical relevance: smoking behavior (*1*), coffee consumption behavior (*2*), alcohol drinking behavior (*3*), and mental disability (*4*) or illness (*5*). In contrast, there has been limited exploration of common complex-behaviors such as aggression (*6*), happiness (*7*) or social phobia (*8*).

Largely due to heterogeneity, most complex traits are difficult to map in humans. Dozens to thousands of variants can each contribute minute amounts to heritable risk, and this can differ dramatically in different ethnicities and their subgroups. Many major breakthroughs in human genetics have resulted from studying isolated populations and multigenerational families. The advantages of those latter approaches, and others, are dramatically exaggerated in dogs: (i) There are approximately 400 dog breeds, each on the order of 100-fold less genetically-complex than the full population. Thus, compared to humans and their major ethnic groups, dogs are much more similar within breeds and much more different across breeds. (ii) Dogs are often part of a family or working environment and receive high levels of health care. Lastly, iii) dogs have more phenotypic variation than any other land mammal; and much of that variation is the result of "evolutionary" selection under domestication. The strengths of dog models of complex genetics have been exploited mainly in the area of cancer (*9*), but recently also in behavior. Examples include investigation of obsessive compulsive disorders in select breeds (*10*, *11*) and diverse behavioral traits in one breed (nerve stability, affability, wariness, adaptability, sharpness, activity and reactions during blood draws (*12*)).

Strikingly, multiple groups recently showed that most genetic variation associated with diverse morphological traits in dogs can be mapped by cross-breed GWA using only breed stereotypes and about a dozen or more individuals each from dozens of diverse breeds (*13-15*). This revealed that a great extent of the genetic variation in domesticated dogs was present prior to breed creation. A good example is the trait of body size, the majority of which is explained by six gene variants in all but very large breeds (see the following and refs within: (*16*)). Those canine genes are relevant to known size-biology across phyla (e.g., *IGF1*, IGF1 receptor, growth hormone receptor, *HMGA2* and *SMAD2*) and others indicate opportunities for discovery (e.g., *STC2*). A landmark GWAS from 2008 reported many loci and five candidate genes associated with the following behavioral traits: herding, boldness, excitability, pointing, and trainability (*13*). A more recent study that used the same stereotype data, and a cohort with different breeds (*15*), did not replicate those findings for one overlapping trait, boldness (which was associated with loci on five chromosomes in the first study and a single other chromosome in the second). The chr10 region that was associated with boldness in the newer study, between the *MSRB3* and *HMGA2* genes, was the same as was strongly associated with two morphological traits – reduced ear erectness and small size. Although each of the three traits appeared to be associated with a different haplotype, with one exception, all bold breeds were erect-eared and small, and vice versa for non-bold breeds. This region spans among the most highly-differentiated markers reported from single-marker F_{ST} analysis and, at 2Mb, it is the second largest of such regions. Similarly, Vaysse et al. showed that sociability (attitude toward unknown humans) maps to the highest F_{ST} region in the genome (2.6Mb, chrX), which was shown by others to be associated with skull shape and large size (*14*). To our knowledge, there are no further claims to resolve the various genetic associations or to suggest biological relevance of those loci to boldness and sociability; they appear to be open questions.

Here we report mapping fear and aggression traits associated with genetic variation shared across diverse breeds. These represent very common and important

canine traits in the behavioral veterinary setting (*17*), and in human public health (*18*). It seems likely to us that our findings will also prove to be relevant to human anxiety disorders and aggression, violence and criminality. Additionally, dog is the only animal that was originally domesticated by humans for almost-purely behavioral traits – and arguably is the only predator to be fully domesticated. Fear, aggression and related traits like tameness have long been thought to be central to the domestication of dogs (*19*), and this is supported by experimental domestication of silver foxes (*20*). Both wild wolves and foxes are typically more fearful and aggressive than their domesticated counterparts; however, some dog breeds have been actively selected for enhanced aggressiveness in certain contexts such as fighting, guarding or vermin control. Our findings show that canine fear and aggression that are directed toward strange humans or other dogs share variation that was present prior to the creation of dog breeds. In another work, we report fine mapping analysis that implicates specific single genes at two loci associated with canine fear and aggression (*21*). Both genes are highly tissue-specific – with highest expression levels in the amygdala and pituitary/hypothalamus, respectively – and suggest strong biological relevance for these traits. One variant is protective and the other increases risk of fear and aggression. We discuss below how variation at these loci may have been selected-for during the process of domestication.

Results

Study design

The present study was designed to test whether breed stereotypes of fear and aggression could be mapped by cross-breed GWA. While this concept has been validated for morphological traits, it has not been for behavioral traits. Success here requires two primary elements: biologically-relevant and robust phenotype data (seemingly likely from studies cited below), and the sharing of behaviorally-associated genetic variation across diverse breeds (which is unknown). We used three unrelated breed-specific resources: one of behavioral phenotypes (*22*) and two of breed-specific genotypes (*14*, *15*). The phenotype dataset is derived from C-BARQ owner questionnaires (*23*). In C-BARQ, fear and aggression comprise five and four subtypes, respectively. All but two of these C-BARQ phenotypes (‘dog rivalry’ and ‘touch sensitivity’) were previously validated using a panel of 200 dogs with prior diagnoses of specific behavior problems (*23*). More recently, other studies have also provided criterion validation by demonstrating associations between these phenotypes and particular training outcomes in working dogs (*24*), and the performance of dogs in various standardized behavioral tests (*25*-*27*). There is currently no alternative phenotype resource that approximates the numbers of breeds and traits represented. The two genotype datasets used here are those used to map cross-breed stereotypes and study population genetics by Vaysse and Ratnakumar et al. (a large European collaboration including the LUPA Consortium, led by Webster; 30 breeds, 175,000 SNP markers) and Boyko et al. (a large USA collaboration led by Bustamante and Ostrander; 45,000 SNPs). The behavioral and genotype datasets overlap for 11 breeds and 29 breeds for Vaysse’s and Boyko’s datasets respectively.

Our study design included the following: i) principal components analysis (PCA) of breed phenotypes and genotypes (C-BARQ data and each genotype dataset); ii) discovery study of GWA mapping of published C-BARQ behavioral data (i.e., that corresponding to the Vaysse subset of the 30 most popular American Kennel Club breeds) with the Vaysse genotypes; iii) confirmation study of the Discovery results using C-BARQ and Boyko datasets; and iv) testing of the internal-consistency of the C-BARQ phenotypes and of the prediction-model performance in a second set of breeds for which phenotype data was not previously published. In separate work, we report fine mapping the two peak regions associated with canine fear and aggression directed to other dogs and human strangers (*21*).

Discovery studies of genetic association with fear and aggression

Assuming sufficient power, the greatest potential limitations of the present GWAS's are latent variables, such as cryptic relatedness and batch effects, and population structure. We mitigated these in several ways, beginning by using many breeds instead of few. We used two genotype datasets that represented different genotyping platforms and cohorts with partially overlapping breed contents as discovery and confirmation datasets. Thus, each dataset has different batch effects, population structure and cryptic relatedness. Lastly, we controlled for relatedness and population structure in the GWAS's by using a centered relatedness matrix correction (i.e., the Genome-wide Efficient Mixed Model Association algorithm or GEMMA (*28*)).

Before initiating mapping studies, we conducted Principal Components Analysis (PCA) of the breed-specific C-BARQ data on fear and aggression together with the genotype data from each of the genotype datasets separately (Fig. 1). Component scores of the breeds evaluated (Figs. 1B/D) showed no relevant deviations across the two genetic datasets. Breeds evaluated in both datasets are distributed similarly in both plots. This preliminary evaluation suggests that, despite having different breed makeups in the two datasets, they show consistent results. Using either genotype dataset the following traits are clustered together apart from the others: stranger-oriented fear, stranger-directed aggression and dog-directed aggression. This suggested to us that the three traits could be genetically-related. A second implication of the PCA results is that owner-directed aggression is most distant, and therefore different, from the three clustered traits mentioned above. In other words, the PCA indicated a testable hypothesis that the three clustered traits would share associated loci, but that owner-directed aggression was associated with other loci (we go on to confirm this).

Our initial discovery study tested for genomewide-significant association for fear and aggression traits from C-BARQ using the genotype dataset of Vaysse et al. (*15*) (Figs. 2, 3). This GWAS involved 175,000 SNP markers genotyped in 150 individuals from 11 breeds. Consistent with the PCA results, stranger-oriented fear, and stranger and dog-directed aggression is predominantly associated with many markers at two loci – chr18:23,260,370 and chrX:105,245,495-105,877,339 (CanFam2 assembly) – whereas owner-directed aggression is distinct. Rather, the latter is associated with previously described small-size variation in the *IGF1* gene on chr15:44Mb and with a single marker at chr34:29Mb (Fig. 2C). Dog-oriented fear, which does not purely cluster with stranger-

directed fear/aggression and dog-directed aggression (but is not distant), is also predominantly-associated with the chr18 and X loci. Touch-sensitivity also has both of those loci, but a marker on chr10:11Mb (affecting *HMGA2*; in this dataset, it is the second strongest cross-breed variant associated with small size after *IGF1* (*15*)) has the second strongest signal and chrX:105Mb is considerably weaker. These latter two findings are consistent with the PCA pattern in Figure 1C, where touch-sensitivity and dog-oriented fear are closest to stranger-oriented fear and stranger- and dog-directed aggression. The other fear traits share one, but not both, of the chr18 and X loci: nonsocial fear is associated with chr18 and the same chr10:11Mb marker associated with touch-sensitivity; and separation-related anxiety is associated with chrX:105Mb and chr10:11Mb (chr18:23Mb is suggestive).

Confirmation of genetic association results in a second cohort

We repeated the discovery studies using the same behavioral data, but with the Boyko genotype dataset (*14*) (Figs. 2, 3; Table 1). The data we used was genotypes of 45,000 markers in 327 individuals from 29 breeds, 11 of which overlapped the Vaysse data used above. Because the same breeds were not used in the two studies, this is not strictly a replication study. However, because we are looking at cross-breed association, the findings in one can confirm those in the other. We also expect the paired GWAS's to mitigate false positive hits that are due to latent variables or population structure in the individual studies.

The results of the confirmation GWAS's generally confirmed associations of chr18 and X with stranger and dog-oriented fear and aggression. Both loci are genomewide-significant in dog-oriented fear, but only chr18 is significant in stranger-oriented fear and aggression (chrX was suggestive in both). Dog-directed aggression has no significant hits, but has suggestive evidence for chr18 among the top ten markers. As in the Vaysse GWA above, owner-directed aggression is most strongly associated with several markers that peak within the *IGF1* gene on chr15:44Mb. Whereas the discovery GWAS of dog rivalry has no hits, the confirmation GWAS shows strong association with *IGF1*; and, unlike in the discovery GWAS, this locus is also significant in dog-oriented fear, separation-related anxiety, and touch-sensitivity. We interpret this as predominantly due to the breed make-up of the two cohorts. At the level of genomewide significance, the most similar results between the two studies are for dog-oriented fear (chr18 and X), separation-related anxiety (chr10 and chrX) and touch sensitivity (chr10 and chr18, with chrX being significant in discovery and almost significant in confirmation).

Prediction model and internal consistency of C-BARQ behavioral data

We next tested the C-BARQ instrument for internal consistency of the behavioral data. Specifically, we wanted to know if the behavioral data is robust as well as qualitatively and quantitatively consistent across breeds. Alternative possibilities include low signal to noise ratios, biases due to human breed-ownership behaviors, and biases related to dog status or behavior that affect participation in C-BARQ study. This analysis does not require one to assume that a trait classified as fear or aggression is necessarily so (or that

those can be sub classified in biologically meaningful ways); it is only to test whether the same classification in one group of breeds is the same as that in another group.

Our most robust discovery and confirmation candidates presented above are those four loci significantly-associated with C-BARQ fear and aggression traits in two independent datasets (from different genotyping platforms applied to different breed cohorts): chr10, chr15, chr18 and X. We used those markers to design a phenotype prediction model that could be applied to the breeds for which C-BARQ behavioral data exists, but has not been published previously. The model was created using our results for the published C-BARQ data for the subset of the top 30 American Kennel Club breeds (*22*) for which we had genotype data. It was implemented on all other breeds for which we have genotypes within Vaysse's dataset; this was done without prior knowledge of the C-BARQ breed stereotypes that were being predicted. We thus predicted fear and aggression for 18 breeds with existing, but previously unpublished, C-BARQ data. We used a model design of multiple linear regressions with a stepwise forward selection method (Fig. S1). The discovery breeds' allele frequencies at each of the four loci were used as primary data. We allowed the selection method to keep only significant predictors by trait. Then we used the beta coefficients estimated for the predictors remaining in the model, along with the allele frequencies of the prediction phase breeds, to estimate their breed-predicted value for that trait.

Across all prediction models (Fig. S1), the hit on chr10 was never significant and was thus never included. Chr15 was removed by the method for all traits except for owner-directed aggression and dog rivalry, for which it was the only significant predictor. This evidence suggests small dog size is highly correlated with these types of aggression as previously suggested (*29*), but it is not a strong predictor of other types of fear and aggression. Chr18 hits remained significant in all other traits, with the exception of owner-directed aggression and dog rivalry. In all models, both chr18 hits (chr18:23260370 and chr18:23298242) were included at the beginning of the selection process but only one or the other remained. These two markers are partially correlated and we assume the detection of one or the other is due to the ancestral haplotype diverging across breeds. The chr18 hits were the only significant predictors for dog-oriented fear, dog-directed aggression, nonsocial fear and touch sensitivity. The chrX hits were significant predictors for stranger-oriented fear, stranger-directed aggression and separation-related anxiety. As with chr18, all three chrX hits detected across all analyses (chrX:105245495, chrX:105770058 and chrX:105877339) were included at the beginning of the selection process, but only one remained for each trait. These markers are very highly correlated and therefore their inclusion in the model can be interchangeable. R-square values for all predictions range from 0.54 to 0.9. We observed that beta coefficients for the chr18 hits are always positive and have values double in size to the beta coefficients for the chrX hit, which is always negative (Fig. S1). Our prediction models suggest that chr18 and X have an additive effect on fear and aggression variability. More precisely, our data suggests that the most common chr18 haplotype across dog breeds is protective from some types of fear/aggression (Fig. 4). In contrast, one of the two common haplotypes on chrX (Fig. 4) increases risk.

To evaluate the performance of our prediction models, we calculated predicted values for all breeds not used in the discovery and confirmation phases and compared them to the values obtained from the raw unreleased C-BARQ data (by calculating the difference between them). To determine whether the performance of our prediction is superior to random chance, we first estimated the chance of randomly making a successful prediction within a 95% confidence interval for all of the observed values. Since the sample sizes for each breed is different, the random chance of making a successful prediction within a 95% confidence is also different. The median chance for making a successful prediction by chance across all breeds and traits is 5.82%, ranging from 1.1% to 30.28%. Next, we calculated the cumulative probability of all random chances. We estimated that 11 out of the 162 values predicted would be expected by chance. Lastly, we evaluated if the difference between the observed and predicted values is inside or outside of the confidence interval, and declared a success for every case where our predicted value was inside. Our total number of successful predictions was 95 out of 164 or 58% which is significantly better than random chance ($p<0.0001$). We propose that this validates the internal consistency of C-BARQ behavioral data across breeds. A success and failure matrix that summarizes our predictions is displayed in Figure 5. Our predictions are not skewed in a single direction across traits and breeds. However, our prediction efficiency is better for some traits and breeds than others. That is consistent with the possibilities that other variants are involved, and that those are not uniformly associated with those traits across many or most breeds.

Haplotype analysis and candidate genes

The associated loci and alleles on chr15 (*IGF1*), chr10 (*HMGA2*) and chr3 (*IGF1R*) are the same as those known to be associated with small size across dog breeds. The corroborated hits on chr18 at 23,260,370 and 23,298,242 map between *GNAT3* and *CD36* (Fig. 6A). The region was reported to be associated with allometric dog height and leg length (*14*), and correspondingly demonstrated to harbor a *FGF4*-retrogene insertion that causes chondrodysplasia in several dog breeds (*30*). In the Vaysse dataset, the fear-increasing chr18 haplotype is only present in Dachshunds and Yorkshire Terriers. Both those breeds carry the chondrodysplasia retrogene insertion. However, while the short-legged Dachshund is fixed for the insertion and always exhibits the trait, the insertion is a common minor allele in the Yorkshire Terrier but it never results in the trait (speculated to be due to its small size (*30*)). The following breeds that have ≥20% allele frequencies of the increased-fear haplotype, but lack common chondrodysplasia: Shar-Pei, Beagle, Cocker Spaniel and Husky (both the Beagle and Cocker Spaniel are known to lack the retrogene-insertion allele (*30*)). These observations suggest that the retrogene insertion occurred in the ancestral increased-fear allele, and that the two traits are only linked in that resulting lineage. The corroborated hit on chrX:105,877,339 is within the 2.6Mb LD block with the strongest F_{ST} signal in the dog genome (*15*). Others have mapped large size and skull shape to this locus, but no clear gene implications have been reported (*14*). The peak fear region has one non-coding RNA and two protein coding genes, with *IGSF1* being the closest (Fig. 6B). In most breeds that carry the reduced-fear allele (9/11 in the Vaysse dataset), that allele is on the same haplotype as the increased-size allele. While

the linkage is not nearly as strong between the alleles for increased fear and sociability, they are also correlated in several breeds. These findings at the four loci are consistent with the PCA analysis that shows at least two different fear and aggression phenomena – one of which is associated with dog size.

Analysis of the peak region on chr18 reveals a very low level of LD (Fig. 6A), indicating that the associated haplotype is very old. In contrast to chr18, the associated region on chrX lies within a 2.6-3.7Mb block of strong LD (Fig. 6B; (*14, 15*)). We find the same increased-fear alleles at the two loci are present at frequencies of 43 and 50%, respectively, in extant wolf populations (*31, 32*). Analysis of relevant breeds using whole genome sequence data (*31, 32*) and DNA copy number variation (CNV) data (*33-35*) appears to rule out that the functional variants are protein coding changes or CNVs. Based only on that and the positional results, we would propose the functional variation affected expression of *GNAT3* or *CD36* on chr18 and *IGSF1* (or the flanking genes *ARHGAP36* and *FIRRE*) on chrX. However, the locus on chrX requires detailed breed-specific haplotype analysis to determine if fine mapping is possible within that 2.6Mb LD block. In the Discussion we mention our separate work that analyzes the evolutionary patterns of variation at the two loci and considers the biological relevance of the two genes implicated by that (*GNAT3* and *IGSF1*).

Discussion

Understanding fear and aggression in dogs is important for canine wellbeing, human public health and to understand the process of dog domestication. It also has great potential to lead to medical translation to related mental disorders in humans. Here we mapped behavioral traits by using breed stereotypes to conduct a series of interbreed genome scans. This approach is well developed for morphological traits, but has not been validated for behaviors. In this work, we validate our findings by conducting separate GWAS's using cohorts comprised of partially overlapping breeds (Table 1). We then provide additional validation by developing a predictive model and applying it successfully to a different group of breeds. This also serves to demonstrate internal consistency of C-BARQ phenotyping across breeds.

Our PCA of breed genotypes and phenotypes indicates that some types of fear and aggression are related to each other (stranger-oriented fear and dog-/stranger-directed aggression), but are distinct from others such as owner-directed aggression. This pattern was mirrored by the results of our behavioral GWAS's which identified two genome loci associated with the former (also shared with dog-oriented fear) and another two associated with the latter (also shared with dog rivalry). Notably, owner-directed aggression and dog rivalry are associated with the same variation in *IGF1* that is known to have the greatest contribution to small-size across dog breeds (the former trait is also associated with the small-size variant at the IGF1 receptor gene). This finding is consistent with previous reports that i) there is a highly-significant correlation between the behaviors of owner-directed aggression and dog rivalry, ii) this correlation is independent of dog- and stranger-directed aggression, and iii) these behaviors are associated with breeds of small to medium size (*29, 36, 37*). And owner personality does not necessarily predispose to owner-directed aggression, which is thus an apparent dog

trait (*38*). Some of those studies also showed a correlation between small size and stranger-oriented fear and aggression, dog-oriented fear, separation anxiety, and touch sensitivity (*29*). That is supported by our finding in the confirmation GWAS's that the same small-dog *IGF1* allele is associated with the latter three traits. It is unclear whether the behavioral associations with small-size gene variants are due to developmental, physiological or psychological effects; all seem probable.

The loci on chr18 and X are particularly interesting because they are novel and associated strongly with the related traits of fear and aggression directed to dogs and human strangers. Variants in those regions are also associated with morphological traits such as leg length and body size, respectively. The closest genes to the chr18 and X association peaks are *GNAT3* (Gustducin alpha, the G protein alpha subunit for bitter, sweet and umami taste cells (*39*)) and *CD36*, and *IGSF1* (which results in a congenital syndrome affecting Thyroid and Growth Hormones when mutant in humans (*40*)), respectively. The fear/aggression peak near *IGSF1* lies within a 2.6-3.7Mb region of strong LD that includes flanking peaks for size and sociability (*14, 15*). In a separate report, we demonstrate how evolutionary footprints can be used to fine map functional variation across dog breeds (*21*). There, we show that the reduced-fear/aggression allele is often in perfect LD with the increased-size allele, but much less so with regards to the increased-sociability peak allele. We also provide evidence that the fear/aggression functional variants on chr18 and X are most likely to affect expression of *GNAT3* and *IGSF1*, respectively. [Notably, *IGSF1* is co-expressed with its two flanking genes, *ARHGAP36* and the long non-coding RNA *FIRRE* (*21, 41*).] Both of those neuronal genes have strong biological relevance at the level of neuroanatomy (see *21 and refs. within*). Outside taste receptor and other chemosensory cells (and a subset of vomeronasal interneurons (*42*)), *GNAT3* is most highly expressed in the amygdala (*21*). *IGSF1* is predominantly and very abundantly expressed in two brain regions - the pituitary and parts of the hypothalamus (*21, 43*). Further studies are necessary to determine if the size and fear/aggression traits on chrX are due to the same, distinct or overlapping variation. It is interesting that at least three genes in the region are co-expressed in tissues involved in determination of body size (the pituitary gland and hypothalamus), and that mutations in *IGSF1* affect human size (*40*). Rat studies have revealed that mRNA expression of both genes are regulated in fear-relevant models: *GNAT3* in amygdala under pain stimuli (*44*); and *IGSF1* in cortex after stress or tactile stimuli (*45*). Thus both genes are strong candidates for fear relevance: they are neuronal and associated with stress/anxiety according to interbreed dog GWA, neuroanatomy and biology.

Because our model of the chr18 and X variation was successful in predicting the relevant fear and aggression behaviors in a third group of non-overlapping breeds, we believe these markers can be used to predict and, in part, explain such behavior across many dog breeds. However, the behavioral stereotypes of some breeds were not explained by our predictive model, and many breeds have not been tested. It seems likely that many breeds have epistatic variation with chr18 and X variants or have other variants that are less commonly associated with these traits across dog breeds. All of these issues can be addressed by studies of individual breeds. Since we have only found the common

co-occurrence of the increased-fear/aggression variation on chr18 and X in small and medium dogs, it will be interesting to see if this is also present in large dogs bred for aggression or fighting.

Dogs were the first animals to be domesticated by humans. New studies indicate that dog domestication has a single origin in southern East Asia ~33,000YA, followed by migration to the Middle East, Africa and Europe ~15,000YA (*46*). Canine fear and aggression are of great interest because those traits – mainly the loss of fear of humans – are widely believed to underlie the mechanism for domestication (*19*). We thus considered whether the variation we identified at chr18 and X could have been involved in the process of dog domestication. Our analysis of published genetic variation from extant wolf populations across the world shows that both fear/aggression alleles (protective on chr18 and risk on chrX) are common in wolves – approximately 50% (*31, 32*) and 43% (*32*), respectively. In domesticated dogs, 27 of 30 breeds in Figure 4 have the reduced fear/aggression chr18 allele at a frequency of at least 75% (19/27 are fixed at a level of 100%). In contrast, only 10/27 breeds have the reduced-fear/aggression chrX allele at frequencies greater than that of wolves (starting at 68% allele frequency; 5 are fixed at a level of at least 95%). The common high-frequencies of the chr18 reduced-fear/aggression allele across dog breeds are consistent with selection of reduced fear and aggression in the domestication of dogs. Interpretation of the chrX region seems less clear because the majority of breeds have the increased-fear/aggression allele. However, in parallel studies we show that the reduced fear/aggression allele has been under selection (i.e., its haplotype size is always large whereas the alternative allele is generally far smaller (*21*)). The high frequency of the four alleles at chr18/X in extant wolf populations may seem counterintuitive in a model where ancestral wolves were generally more fearful and aggressive than domesticated dogs, but this could be explained by balancing selection or recent positive selection.

A concrete understanding of dog domestication is expected soon as there is a major effort underway to sequence ancient dog genomes (*47*). Our prediction is that positive selection of the reduced-fear chr18 and chrX variants was part of the domestication process. We further propose that humans have obscured those roots in modern dog breeds by selecting for increased aggression or for other linked morphological or behavioral traits (e.g., for short legs at chr18 or for increased aggression or smaller size at chrX). We expect there are many other fear/aggression-associated loci that are more difficult to map across breeds. But it seems likely that having two very common alleles influencing these behaviors led to frequent selection between those to set a level of reactivity and disposition towards dogs and humans. In this regard, it seems plausible that the correlation between small size and the highest levels of fear/aggression is because the same behavior in large dogs is generally unacceptable to humans (*36*). In our principal mapping study, only small dogs – mainly Dachshund and Yorkshire Terrier – have high frequency of fear/aggression associated alleles at both chr18 and X. The fact that dog and stranger oriented fear and aggression are generally much more strongly associated with chr18 and X variants than with *IGF1*/chr15, *HMGA2*/chr10 and *IGF1R*/chr3 small-size variants further establishes that small size is not the predominant

cause. However, it is also notable that a subset of large breeds carries the chrX reduced-fear/aggression and increased-size variants in perfect LD.

The biochemistry and neuroanatomy of the emotions of fear and aggression are highly conserved in vertebrates, and some argue this is true across the animal kingdom (*48*). Across vertebrates the most immediate response to extreme threat involves the transmission of different sensory signals through the following sequence of brain regions (referred to as the low road): thalamus, amydgala, hypothalamus and pituitary gland, which sends nerve and hormonal signals to the adrenal glands, which in turn direct acute (through noradrenaline/adrenaline) and sustained (through glucocorticoid hormones such as cortisol) stress responses. This low road corresponds to innately programmed responses and is associated with emotions such as fear and anger. A parallel cognitive pathway that is referred to as the high road diverges at the thalamus, by then going to primary sensory and association centers in the cortex before continuing to the amygdala. Both pathways also involve bidirectional signaling with the hippocampus. Thus, while the immediate response to fear may be predominantly innate and emotional, it is not completely separate from cognition. There is extensive molecular and behavioral evidence that the hypothalamic-pituitary-adrenal (HPA) axis is the most critical driver of behavioral stress. Biochemical pathways implicated in social fear and aggression include signaling by serotonin and dopamine, and neuropeptides such as the predominantly-hypothalamic oxytocin and vasopressin (*26*). Notably, the domestication of another canid, the fox, resulted in foxes with greatly reduced HPA activity (*20*). After 45 generations of selection for tameness in foxes, basal blood cortisol levels were reduced three-fold and stress-induced levels five-fold (compared to normal foxes). Domesticated foxes also have increased levels of brain serotonin, consistent with its inhibitory effect on aggression. New analysis of selective sweep regions associated with domestication of pigs showed that *GNAT3* lies in one of the sweep regions of European (but not Asian) pigs (*49*); but it remains to be seen if *GNAT3* variation is directly associated with that domestication event. Our findings that loci spanning *GNAT3* (which is highly expressed in the amygdala) and *IGSF1* (highly expressed in the hypothalamus and pituitary gland) are associated with canine fear and aggression are thus consistent with a very large body of work implicating the HPA axis.

It is clear from animal and human studies that fear and aggression are often associated, but it is not always in the same direction (*50*). Based on human behavior and pharmacology, the links between anxiety and aggression are very complex. Similarly, early life stresses in people and animals are associated in complex ways with anxiety disorders and aggression in adulthood. Early life stress generally involves changes in the HPA axis, and results in increased anxiety and altered social and aggressive behaviors. Animal models with a profile similar to the dog case presented here – in which both fear and aggression are elevated – are rare. Examples of knockout mouse models that have this property include those for enkephalin and α-calcium-calmodulin kinase II (*51, 52*). Selective breeding has also yielded strains that have increases in both anxiety and aggression. One of those is the North Carolina mouse, in which acute diazepam treatment reduces both anxiety and aggression (*53*). The other example is a strain of Novosibirsk Norway rats that was bred for increased aggression to humans (*54*). It is not immediately

clear which human conditions may be most relevant to the present dog model at chr18 and X. Most likely those will include anxiety disorders. Notably, some anxiety conditions are associated with increased aggression, and this includes a subset of those affected by social anxiety.

Now that we have identified common variants associated with fear and aggression across dog breeds, the segregation of these variants within breeds provide powerful models in which to dissect their biology at the levels of neuroanatomy, physiology and behavior. Among the areas of research on fear and aggression (*48*) that are ongoing in dogs and will be vastly accelerated by genetic handles are i) development and environmental-malleability (*27, 55*); ii) molecular/biochemical and imaging descriptions at baseline and under acute stress (*20*), iii) effects on mental and physiological states in the life course (*24-26, 36*), and iv) feasibility to mitigate negative effects through cognitive or pharmacological treatments (*56, 57*). In parallel, it will be important to determine the molecular mechanisms of these fear/aggression variants, and to identify their interactions with other genes and environmental factors.

Materials and Methods

Experimental design overview

Since the analysis of dog breed data is highly vulnerable to population structure issues and false positive detection, we designed our analysis in two phases: first, a discovery phase, where available SNP data was used to map aggression and fear behavioral traits. We designed this phase to analyze two independent SNP datasets using two independent behavioral phenotypes for aggression and fear. Significant hits were taken then into a second phase for validation. To validate our findings we evaluated the performance of behavioral values predicted for breeds not included in the discovery phase. The expectation was to be able to predict behavioral traits from a few markers.

Discovery Behavioral phenotypes

Phenotype values for aggression and fear variables were taken directly from the C-BARQ values previously published for the top 30 most popular breeds of the AKC (*22*). We refer to this collection of behavioral phenotypes as "C-BARQ phenotypes". This C-BARQ dataset is a collection of owner reported behavioral data of AKC registered dogs. Only the breeds for which SNP data (see SNP datasets section) were available were included in the analysis; therefore a total of 6,818 subjects were used to determine the phenotypical values. C-BARQ data decomposes aggression into 4 classes: Stranger-directed aggression (towards unfamiliar humans), Dog-directed aggression (towards unfamiliar dogs), Owner-directed aggression and Dog rivalry (towards familiar humans and dogs, respectively). In a similar way C-BARQ data decomposes fear into 5 classes: Stranger-oriented fear (towards unfamiliar humans), Dog-oriented fear (towards unfamiliar dogs), Nonsocial fear (towards environmental phenomena), Separation-related anxiety (being left alone by the owner) and Touch sensitivity.

Validation Behavioral phenotypes

To validate the hits mapped using the discovery data, we inferred predictions based on the markers detected for the breeds from the Vaysse dataset (*15*) that were not included in the top 30 most popular breeds (*22*). Our predicted C-BARQ values for 18 dog breeds (see phenotype prediction analysis section) were compared to C-BARQ phenotypes calculated from unreleased raw data (Serpell, unpublished). These C-BARQ phenotypical values were obtained from 2,130 subjects of 18 breeds. Only one breed (Greenland Sledge dog) included in the Vaysse dataset had no data available in the C-BARQ database and thus was excluded from the prediction analysis.

SNP datasets

Two previously published SNP datasets were used in this study. The first dataset contained ~175,000 SNPs on the Illumina CanineHD array; we refer to this dataset as the "Vaysse dataset" (*15*). The second dataset contained ~45,000 SNPs on the Affymetrix v.2 Canine array; we refer to this dataset as the "Boyko dataset" (*14*). The Vaysse dataset contained a total of 456 subjects representing 30 dog breeds while the Boyko dataset contained 890 subjects representing 80 dog breeds. Since the stereotypic phenotype data was not available for all the breeds included in each of the datasets, only those for which phenotypes were available were kept; therefore, the Vaysse dataset contained 150 subjects from 11 dog breeds while the Boyko dataset contained 327 subjects from 29 dog breeds. All dog breeds included in the Vaysse dataset used were also included in the Boyko dataset. Since the Vaysse dataset has higher resolution and cleaner signal as compared to the Boyko dataset (see original publications for more details) we designated the Vaysse dataset as our main discovery dataset while the Boyko dataset would be used for further validation of the findings detected on the Vaysse dataset. Both datasets are independent of each other and no subjects are shared between them.

Genomewide Association Analysis and Mapping

The preparation of datasets and subject removal were carried out in PLINK v1.07 (*58*). Principal component analysis evaluation was performed on GCTA v.1.24 (*59*). All association analysis were performed on GEMMA v.0.94.1 (*28*). Population structure was removed by using the centered relatedness matrix correction; the association tests were performed using the Univariate Linear Mixed Model using the Likelihood Ratio Test. Genomewide significance was declared for the Vaysse dataset for a *P*-value equal or less than 1×10^{-8}; for the Boyko dataset, genomewide significance was declared for a *P*-value equal or less than 1×10^{-5}. GEMMA was run on the Ohio Supercomputer Center's Oakley Cluster (www.osc.edu) for faster processing. To avoid irreproducibility issues, no dataset trimming or LD clustering were performed on the SNP data. All Manhattan plots were generated by SAS v.9.3 from GEMMA outputs. Genomewide significant hits were mapped on the UCSC Genome Browser (*60*) but coordinates were lifted to take advantage of the enhanced annotation available from the Broad Institute CanFam3 Improved Annotation Data V.1 (*61*) since the original SNP coordinates provided were CanFam2.

Haplotype analysis of genomewide significant hits

LD blocks of adjacent top significant hits and haplotype determination were evaluated by Haploview v.4.2 (*62*). Only top hits were included in the haplotype determination. For hits within the X chromosome all subjects were deemed as females since the Vaysse dataset is not annotated by sex.

Phenotype Prediction Analysis

Significant hits detected across the Vaysse and Boyko datasets for each trait in the discovery phase were further evaluated for their predictive performance. For this part, allele frequencies for the top significant hit for all dog breeds used in the discovery phase and dog breeds not included in the discovery phase (prediction phase breeds) were calculated using PLINK v1.07. Each of the significant markers' allele frequencies were linearly regressed using a stepwise forward selection method based on an inclusion/exclusion alpha cutoff of ≤ 0.05 excluding the intercept. All statistical modeling was performed on SAS v9.3. For each iteration, the selection process tested if any of the allele frequency markers would significantly contribute to the model in terms of P-value significance; if the marker became non-significant after the addition of another marker to the model then it was removed from the model. The final iteration contained all markers that were below the alpha inclusion/exclusion criteria. Once the equation parameters were estimated, predictions were made using the beta coefficients for each marker that remained in the model with the allele frequencies obtained from the prediction phase breeds. By this method, a prediction value for each of the traits for every breed evaluated was generated.

To evaluate the performance of the aggression and fear predictions made, we calculated mean C-BARQ values observed from the unreleased raw database and calculated 95% confidence intervals for each. Sample size and standard errors used to estimate each of the confidence intervals corresponded to the sample size/standard errors of each breed from the unreleased raw dataset for the respective trait. Since the C-BARQ values are contained within a fixed scale that goes from 0 to 4, a uniform distribution was assumed to calculate the cumulative probability associated to each confidence interval range. With this, it was possible to estimate the chance of making a correct prediction by a random guess. We calculated the difference between the observed and our predicted value and, if the difference was contained within the confidence interval of the observed value we considered the prediction a success, while in the case where the predicted value was outside of the confidence interval, it was considered a failure. To determine if the performance of our prediction outcome was superior to a random guess outcome, we tested using a one-tailed Fishers Exact test if the success /failure counts of our prediction outcome were superior to the random guess outcome. Since our models were generated independent of each other, we did not explore the prediction efficiency scenario where a trait is evaluated given that another trait or traits estimates are known. In summary, we evaluated if the chance of correctly predicting the set of values by random guessing was comparable to the predictions made using our models based on the genomewide significant markers for each trait.

References and Notes

1. T. E. Thorgeirsson *et al.*, A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638-U639 (2008).
2. M. C. Cornelis *et al.*, Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Molecular Psychiatry* **20**, 647-656 (2015).
3. L. J. Bierut *et al.*, A genome-wide association study of alcohol dependence. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 5082-5087 (2010).
4. L. Vissers *et al.*, A de novo paradigm for mental retardation. *Nature Genet.* **42**, 1109-+ (2010).
5. R. Uher, The role of genetic variation in the causation of mental illness: an evolution-informed framework. *Molecular Psychiatry* **14**, 1072-1082 (2009).
6. I. W. Craig, K. E. Halton, Genetics of human aggressive behaviour. *Hum Genet* **126**, 101-113 (2009).
7. C. A. Rietveld *et al.*, Molecular genetics and subjective well-being. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 9692-9697 (2013).
8. J. Gelernter, G. P. Page, M. B. Stein, S. W. Woods, Genome-wide linkage scan for loci predisposing to social phobia: Evidence for a chromosome 16 risk locus. *American Journal of Psychiatry* **161**, 59-66 (2004).
9. C. E. Alvarez, Naturally Occurring Cancers in Dogs: Insights for Translational Genetics and Medicine. *ILAR Journal* **55**, 16-45 (2014).
10. R. Tang *et al.*, Candidate genes and functional noncoding variants identified in a canine model of obsessive-compulsive disorder. *Genome Biology* **15**, R25 (2014).
11. K. Tiira *et al.*, Environmental Effects on Compulsive Tail Chasing in Dogs. *PLoS ONE* **7**, e41684 (2012).
12. J. H. Ha, M. Alam, D. H. Lee, J. J. Kim, Whole Genome Association Study to Detect Single Nucleotide Polymorphisms for Behavior in Sapsaree Dog (Canis familiaris). *Asian Australas. J. Anim. Sci* **28**, 936-942 (2015).
13. P. Jones *et al.*, Single-Nucleotide-Polymorphism-Based Association Mapping of Dog Stereotypes. *Genetics* **179**, 1033-1044 (2008).
14. A. R. Boyko *et al.*, A Simple Genetic Architecture Underlies Morphological Variation in Dogs. *PLoS Biol* **8**, e1000451 (2010).
15. A. Vaysse *et al.*, Identification of Genomic Regions Associated with Phenotypic Variation between Dog Breeds using Selection Mapping. *PLoS Genet* **7**, e1002316 (2011).
16. M. Rimbault *et al.*, Derived variants at six genes explain nearly half of size reduction in dog breeds. *Genome Res.* **23**, 1985-1995 (2013).
17. J. A. Serpell, Y. Y. Hsu, Effects of breed, sex, and neuter status on trainability in dogs. *Anthrozoos* **18**, 196-207 (2005).
18. K. L. Overall, M. Love, Dog bites to humans - demography, epidemiology, injury, and risk. *Journal of the American Veterinary Medical Association* **218**, 1923-1934 (2001).
19. R. Coppinger, R. Schneider, in *The domestic dog : its evolution, behaviour, and interactions with people,* J. A. Serpell, Ed. (Cambridge University Press, Cambridge, New York, 1995), chap. 21-47.
20. A. V. Kukekova, S. V. Temnykh, J. L. Johnson, L. N. Trut, G. M. Acland, Genetics of behavior in the silver fox. *Mamm Genome* **23**, 164-177 (2012).

21. I. Zapata, C. E. Alvarez, Fine-mapping and biological relevance implicate *GNAT3* and *IGSF1* in canine fear and aggression. (manuscript submitted for publication).
22. J. Serpell, D. Duffy, in *Domestic Dog Cognition and Behavior,* A. Horowitz, Ed. (Springer Berlin Heidelberg, 2014), chap. 2, pp. 31-57.
23. Y. Y. Hsu, J. A. Serpell, Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *Journal of the American Veterinary Medical Association* **223**, 1293-+ (2003).
24. D. L. Duffy, J. A. Serpell, Predictive validity of a method for evaluating temperament in young guide and service dogs. *Applied Animal Behaviour Science* **138**, 99-109 (2012).
25. S. Barnard, C. Siracusa, I. Reisner, P. Valsecchi, J. A. Serpell, Validity of model devices used to assess canine temperament in behavioral tests. *Applied Animal Behaviour Science* **138**, 79-87 (2012).
26. R. H. De Meester *et al.*, A preliminary study on the use of the Socially Acceptable Behavior test as a test for shyness/confidence in the temperament of dogs. *Journal of Veterinary Behavior: Clinical Applications and Research* **3**, 161-170 (2008).
27. P. Foyer, N. Bjällerhag, E. Wilsson, P. Jensen, Behaviour and experiences of dogs during the first year of life predict the outcome in a later temperament test. *Applied Animal Behaviour Science* **155**, 93-100 (2014).
28. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* **44**, 821-824 (2012).
29. P. D. McGreevy *et al.*, Dog Behavior Co-Varies with Height, Bodyweight and Skull Shape. *PLoS ONE* **8**, e80529 (2013).
30. H. G. Parker *et al.*, An Expressed Fgf4 Retrogene Is Associated with Breed-Defining Chondrodysplasia in Domestic Dogs. *Science* **325**, 995-998 (2009).
31. E. Axelsson *et al.*, The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* **495**, 360-364 (2013).
32. B. Bai *et al.*, DoGSD: the dog and wolf genome SNP database. *Nucleic Acids Research*, (2014).
33. J. Berglund *et al.*, Novel origins of copy number variation in the dog genome. *Genome Biol* **13**, R73 (2012).
34. W. K. Chen, J. D. Swartz, L. J. Rush, C. E. Alvarez, Mapping DNA structural variation in dogs. *Genome Res* **19**, 500-509 (2009).
35. T. J. Nicholas, C. Baker, E. E. Eichler, J. M. Akey, A high-resolution integrated map of copy number polymorphisms within and between breeds of the modern domesticated dog. *BMC Genomics* **12**, 414 (2011).
36. D. L. Duffy, Y. Hsu, J. A. Serpell, Breed differences in canine aggression. *Applied Animal Behaviour Science* **114**, 441-460 (2008).
37. D. L. Duffy, K. A. Kruger, J. A. Serpell, Evaluation of a behavioral assessment tool for dogs relinquished to shelters. *Preventive Veterinary Medicine* **117**, 601-609 (2014).
38. N. H. Dodman, R. Moon, M. Zelin, Influence of owner personality type on expression and treatment outcome of dominance aggression in dogs. *J Am Vet Med Assoc* **209**, 1107-1109 (1996).
39. L. Ruiz-Avila, G. T. Wong, S. Damak, R. F. Margolskee, Dominant loss of responsiveness to sweet and bitter compounds caused by a single mutation in alpha-gustducin.

Proceedings of the National Academy of Sciences of the United States of America **98**, 8868-8873 (2001).
40. Y. Sun *et al.*, Loss-of-function mutations in IGSF1 cause an X-linked syndrome of central hypothyroidism and testicular enlargement. *Nature Genet.* **44**, 1375-1381 (2012).
41. C. Wu, I. MacLeod, A. I. Su, BioGPS and MyGene.info: organizing online, gene-centric information. *Nucleic Acids Research* **41**, D561-D565 (2013).
42. T. Ogura, K. Krosnowski, L. Zhang, M. Bekkerman, W. Lin, Chemoreception Regulates Chemical Access to Mouse Vomeronasal Organ: Role of Solitary Chemosensory Cells. *PLoS ONE* **5**, e11924 (2010).
43. E. S. Lein *et al.*, Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168-176 (2007).
44. I. J. Rampil, D. H. Moller, A. H. Bell, Isoflurane modulates genomic expression in rat amygdala. *Anesth. Analg.* **102**, 1431-1438 (2006).
45. F. C. Zucchi *et al.*, Lifetime stress cumulatively programs brain transcriptome and impedes stroke recovery: benefit of sensory stimulation. *PLoS One* **9**, e92130 (2014).
46. G. D. Wang *et al.*, Out of southern East Asia: the natural history of domestic dogs across the world. *Cell Res*, (2015).
47. D. Grimm, Dawn of the dog. *Science* **348**, 274-279 (2015).
48. D. J. Anderson, R. Adolphs, A framework for studying emotions across species. *Cell* **157**, 187-200 (2014).
49. N. B. Sutter *et al.*, Extensive and breed-specific linkage disequilibrium in Canis familiaris. *Genome Res.* **14**, 2388-2396 (2004).
50. E. K. Karlsson *et al.*, Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat Genet* **39**, 1321-1328 (2007).
51. G. Larson *et al.*, Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proceedings of the National Academy of Sciences* **109**, 8878-8883 (2012).
52. J. M. Smith, J. Haigh, THE HITCHHIKING EFFECT OF A FAVORABLE GENE. *Genetical Research* **23**, 23-35 (1974).
53. G. Ru, B. Terracini, L. T. Glickman, Host related risk factors for canine osteosarcoma. *Vet. J.* **156**, 31-39 (1998).
54. B. Charlesworth, A hitch-hiking guide to the genome: a commentary on 'The hitch-hiking effect of a favourable gene' by John Maynard Smith and John Haigh. *Genetics Research* **89**, 389-390 (2007).
55. M. Morrow *et al.*, Breed-dependent differences in the onset of fear-related avoidance behavior in puppies. *Journal of Veterinary Behavior: Clinical Applications and Research* **10**, 286–294 (2015).
56. N. H. Dodman *et al.*, Use of fluoxetine to treat dominance aggression in dogs. *J Am Vet Med Assoc* **209**, 1585-1587 (1996).
57. Y. Uchida, N. Dodman, J. DeNapoli, L. Aronson, Characterization and treatment of 20 canine dominance aggression cases. *J Vet Med Sci* **59**, 397-399 (1997).
58. S. Purcell *et al.*, PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American journal of human genetics* **81**, 559-575 (2007).
59. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics* **88**, 76-82 (2011).

60. W. J. Kent *et al.*, The Human Genome Browser at UCSC. *Genome Res.* **12**, 996-1006 (2002).
61. M. P. Hoeppner *et al.*, An Improved Canine Genome and a Comprehensive Catalogue of Coding Genes and Non-Coding Transcripts. *PLoS ONE* **9**, e91172 (2014).
62. J. C. Barrett, B. Fry, J. Maller, M. J. Daly, Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265 (2005).

Acknowledgments:
We are grateful to Drs. Vaysse, Ratnakumar and Webster and their co-authors and members of the LUPA Consortium, and Drs. Boyko, Bustamante and Ostrander, and their co-authors, for the availability of the rich genotype data from the Vaysee et al. and Boyko et al. studies cited here. We thank the dog owners who contributed to the C-BARQ database of behavioral phenotypes. C.E.A. is supported by Department of Defense CDMRP grant award no. W81XWH-11-2-0224, American Kennel Club CHF award no. 01660 and Morris Animal Foundation award no. D13CA-073. The authors declare that they have no conflicts of interest. The data used to conduct the present studies were previously published and are referenced in the text.

Fig. 1: PCA analysis of behavioral traits. Since the two GWA datasets only partially overlapped in breed content, PCA analysis was performed to evaluate if the breed makeup affects the distribution structure of the variables. (A, B) Component pattern and component scores, respectively, of C-BARQ behavioral traits on the matching breeds in the Boyko dataset. (C, D) Component pattern and component scores of C-BARQ behavioral traits on the matching breeds in the Vaysse dataset.

Fig. 2: Manhattan plots of C-BARQ aggression traits. Vertical lines indicate relevant and consistent hits across the two GWA datasets. In each panel, the top plot corresponds to the Vaysse dataset while the bottom plot corresponds to Boyko dataset. (A) Stranger-directed aggression. (B) Dog-directed aggression. (C) Owner-directed aggression. (D) Dog rivalry.

Fig. 3: Manhattan plots of C-BARQ fear traits. Vertical lines indicate relevant and consistent hits across the two GWA datasets. In each panel, the top plot corresponds to Vaysse dataset while the bottom plot corresponds to Boyko dataset. (A) Stranger-oriented fear. (B) Dog-oriented fear. (C) Nonsocial-oriented fear. (D) Separation-related anxiety. (E) Touch-sensitivity.

Fig. 4: Haplotype distribution across dog breeds based on novel chr18 and X markers associated with aggression and fear. Haplotypes were defined only on the alleles of the top markers detected in this study. Allele distributions are color coded on a

gradient that goes from yellow (fixed for allele 1) to blue (fixed for allele 2). Allele label letters are arbitrary.

Fig. 5: Success/failure matrix of predicted values. Green fill indicates a successful prediction and red is a failed prediction. Columns correspond to aggression and fear traits while rows correspond to dog breeds predicted. Totals are the sum of successful prediction within the column/row; columns and rows are sorted in numerical order and cells have a fill color gradient that goes from red (worse) to green (best).

Fig. 6: Linkage Disequilibrium plots for chr18/X fear and aggression loci. (A) Chr18 GWA locus for fear/aggression (using C-BARQ phenotypes and the Vaysse genotype dataset). (B) ChrX locus for the same GWAS as A. Genomic coordinates are converted from CanFam2 to CanFam3, and gene information is from the Broad Institute's CanFam3 Improved Annotation Data v.1. Linkage disequilibrium plot was created with Haploview v.4.2.

Supplementary Materials:
Table S1. C-BARQ fear and aggression trait descriptions
Table S2. C-BARQ behavioral phenotypes used to test predictive model
Fig. S1. Summary of prediction equation

Locus (chr:Mb position)	**2:5.7**	**3:43.4**	**3:44.8**	**5:17.2**	**7:5.9**	**10:10.8 -11.1**	**10:43.3**	**11:24.8**	**12:3.2**	**15:44.2**	**18:23.2**	**24:5.9**	**32:8.7**	**34:29.4**	**X:105.2 -106.1**
Stranger-directed fear	V	V							B		V, B				V[3]
Dog-directed fear	V									B	V, B				V[3], B
Stranger-directed aggression				V							V, B				V[1]
Dog-directed aggression											V				V[2]
Touch-sensitivity			B			V, B[6]			B	B	V[5], B				V[3]
Nonsocial fear						V		B	B		V[5], B				V[4]
Separation anxiety						V, B[6]			B	B					V[3]
Dog rivalry							B			B			B		
Owner-directed aggression			B		B					V[7], B		B		V	
Candidate (favored in bold)	***TRDN***	***CERS3***	***IGF1R***	***PVRL1***	***NR5A2***	*MSRB3,* ***HMGA2***	***TMEM182***	***FSTL4***	*ARG1,* ***MED23***	***IGF1***	***GNAT3****, CD36*	***RALGAPA2***	***RASGEF1B***	***SMC4***[8]	*ARHGAP36,* ***IGSF1****, FIRRE, STK26*[9]

Table 1. Summary of canine fear and aggression GWAS results

Fear and aggression trait peaks are given for separate GWAS studies using Vaysse (marked "V"; Illumina HD) and Boyko ("B"; Affymetrix v.2) genotype datasets. Loci shared with both are black and others are gray. Coordinates use CanFam2 assembly.

[1-4]The peak SNPs chrX:105,245,495, chrX:105,770,058, chrX:105,877,339, and chrX:106,189,665 (numbered [1-4] in superscript, respectively) lie within one LD block. At least SNPs 2 and 3 are presumed to implicate the same haplotype/functional variant; candidate genes refer to these peaks.

[5]The peak SNP is 23,298,242 (vs. chr18:23,260,370 for the others).

[6]The peak SNP for Vaysse is chr10:11,169,956 and for Boyko is chr10:10,859,628.

[7]Vaysse peak SNP chr15:44,258,017; Boyko peak SNP chr15:44,226,659

[8]Peak SNP is a coding variant at a generally mammalian-conserved position.

[9]*ARHGAP36*, *IGSF1* and long non-coding RNA *FIRRE* are co-expressed, including in the pituitary gland and hypothalamus (see text).

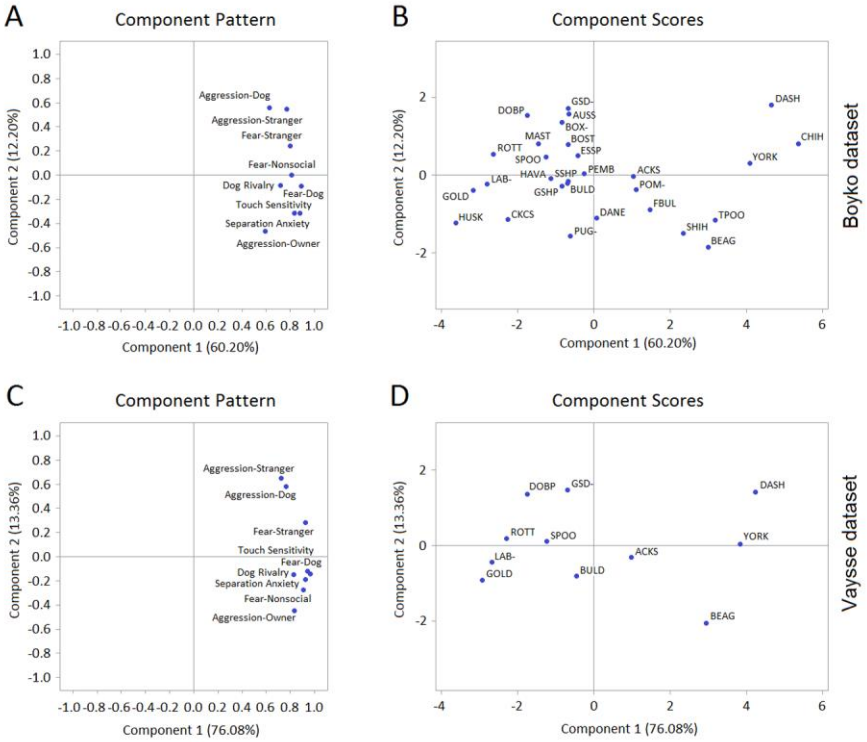

A
Component Pattern
Component 2 (12.20%)
Component 1 (60.20%)
Aggression-Dog
Aggression-Stranger
Fear-Stranger
Fear-Nonsocial
Dog Rivalry
Fear-Dog
Touch Sensitivity
Separation Anxiety
Aggression-Owner
B
Component Scores
Component 2 (12.20%)
Component 1 (60.20%)
Boyko dataset
C
Component Pattern
Component 2 (13.36%)
Component 1 (76.08%)
D
Component Scores
Component 2 (13.36%)
Component 1 (76.08%)
Vaysse dataset

Fig. 1

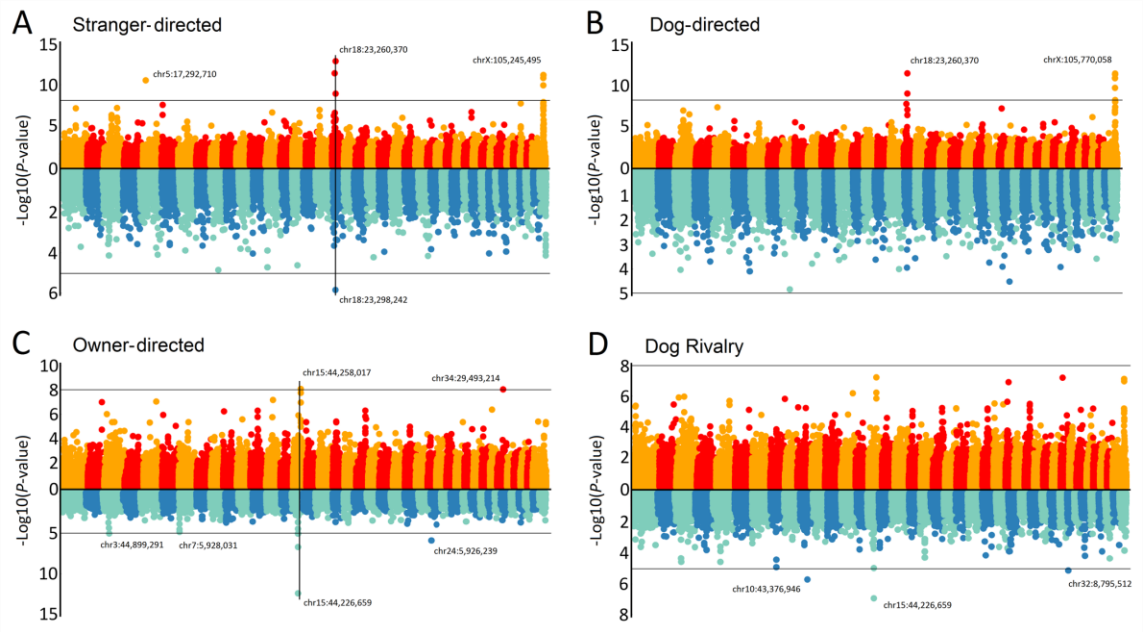

A
Stranger-directed
-Log10(P-value)
chr18:23,260,370
chrX:105,245,495
chr5:17,292,710
chr18:23,298,242
B
Dog-directed
-Log10(P-value)
chr18:23,260,370
chrX:105,770,058
C
Owner-directed
-Log10(P-value)
chr15:44,258,017
chr34:29,493,214
chr3:44,899,291
chr7:5,928,031
chr24:5,926,239
chr15:44,226,659
D
Dog Rivalry
-Log10(P-value)
chr10:43,376,946
chr15:44,226,659
chr32:8,795,512

Fig. 2

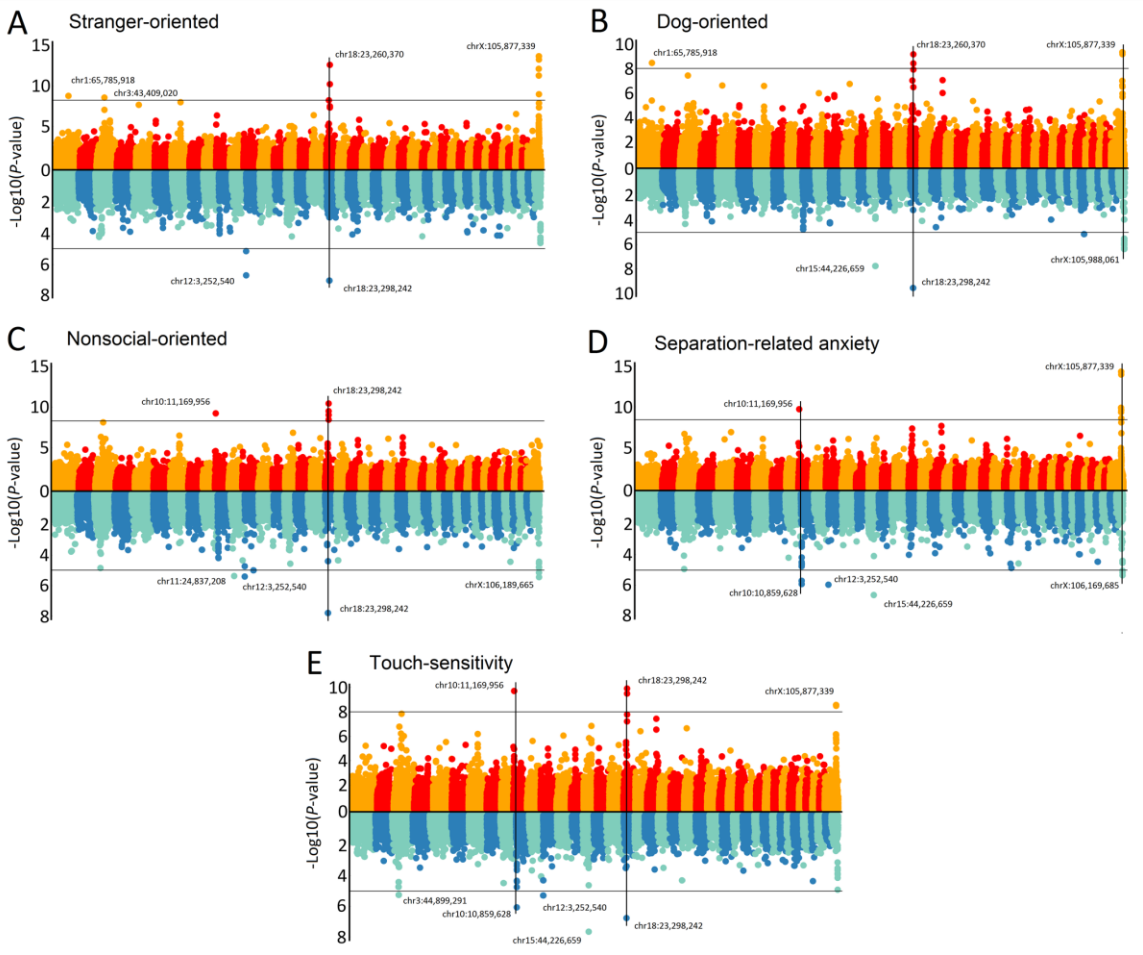

A
Stranger-oriented
-Log10(P-value)
chr1:65,785,918
chr3:43,409,020
chr18:23,260,370
chrX:105,877,339
chr12:3,252,540
chr18:23,298,242
B
Dog-oriented
chr1:65,785,918
chr18:23,260,370
chrX:105,877,339
chr15:44,226,659
chrX:105,988,061
chr18:23,298,242
C
Nonsocial-oriented
chr10:11,169,956
chr18:23,298,242
chr11:24,837,208
chr12:3,252,540
chrX:106,189,665
chr18:23,298,242
D
Separation-related anxiety
chrX:105,877,339
chr10:11,169,956
chr10:10,859,628
chr12:3,252,540
chr15:44,226,659
chrX:106,169,685
E
Touch-sensitivity
chr10:11,169,956
chr18:23,298,242
chrX:105,877,339
chr3:44,899,291
chr10:10,859,628
chr12:3,252,540
chr15:44,226,659
chr18:23,298,242

Fig. 3

	Chr 18: GNAT3 locus				Chr X: IGSF locus			
SNP positions	23260370, 23298242				105245495, 105770058, 105877339			
alleles	CG	CT	AG	AT	GCA	GGA	ACA	AGG
allele label	A	B	C	D	A	B	C	D
Beagle	0	0	0.2	0.8	1	0	0	0
Belgian Tervuren	0	0	0	1	0.958	0	0	0.042
Bernese Mountain Dog	0	0	0	1	0.042	0	0	0.958
Border Terrier	0	0	0	1	0.98	0	0	0.02
Border Collie	0	0	0.125	0.875	1	0	0	0
Brittany Spaniel	0	0	0	1	1	0	0	0
Cocker Spaniel	0	0	0	1	0.964	0	0	0.036
Dachshund	0.667	0	0	0.3333	1	0	0	0
Doberman Pinscher	0	0	0.02	0.98	1	0	0	0
English Bulldog	0	0	0	1	0	0	0	1
English Setter	0	0	0.1667	0.8333	0.917	0	0	0.083
Eurasian	0	0.208	0	0.792	0.167	0	0	0.833
Finnish Spitz	0	0	0	1	0.917	0	0	0.083
German Shepherd	0	0	0	1	0.917	0	0	0.083
Golden Retriever	0	0	0	1	0	0	0	1
Gordon Setter	0.02	0	0	0.98	0.8	0	0	0.2
Greenland Sledge Dog	0	0.042	0	0.958	1	0	0	0
Greyhound	0	0	0	1	1	0	0	0
Irish Wolfhound	0	0	0	1	0	0	0	1
Jack Russell Terrier	0.083	0	0.167	0.75	0.708	0.042	0.042	0.208
Labrador Retriever	0	0	0	1	0.179	0	0	0.821
Newfoundland	0	0	0	1	0	0	0.18	0.82
Norwegian Elkhound	0	0	0	1	0.167	0	0	0.833
Nova Scotia Duck Tolling Retriever	0	0	0	1	0.935	0	0	0.065
Rottweiler	0	0	0	1	0	0	0	1
Schipperke	0	0	0	1	1	0	0	0
Shar Pei	0.136	0.318	0	0.546	0.318	0	0	0.682
Standard Poodle	0	0	0	1	1	0	0	0
Weimaraner	0	0	0	1	0.865	0	0	0.135
Yorkshire Terrier	0.542	0	0.25	0.208	0.958	0	0	0.042
Wolf (Axelsson et al. 2013)	0.286	0.214	0	0.500	N/A	N/A	N/A	N/A
Wolf (Bai et al. 2014)	0	0.500	0	0.500	0.429	0.429	0	0.142

Fig. 4

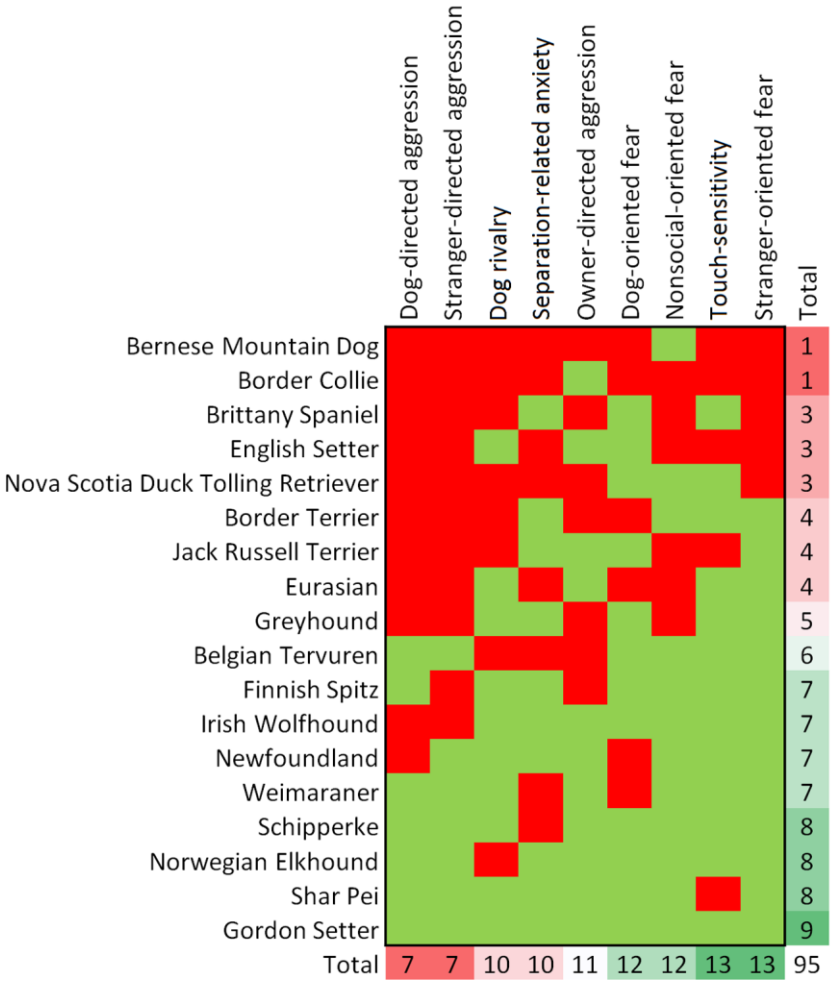

Dog-directed aggression
Stranger-directed aggression
Dog rivalry
Separation-related anxiety
Owner-directed aggression
Dog-oriented fear
Nonsocial-oriented fear
Touch-sensitivity
Stranger-oriented fear
Total
Bernese Mountain Dog 1
Border Collie 1
Brittany Spaniel 3
English Setter 3
Nova Scotia Duck Tolling Retriever 3
Border Terrier 4
Jack Russell Terrier 4
Eurasian 4
Greyhound 5
Belgian Tervuren 6
Finnish Spitz 7
Irish Wolfhound 7
Newfoundland 7
Weimaraner 7
Schipperke 8
Norwegian Elkhound 8
Shar Pei 8
Gordon Setter 9
Total 7 7 10 10 11 12 12 13 13 95

Fig. 5

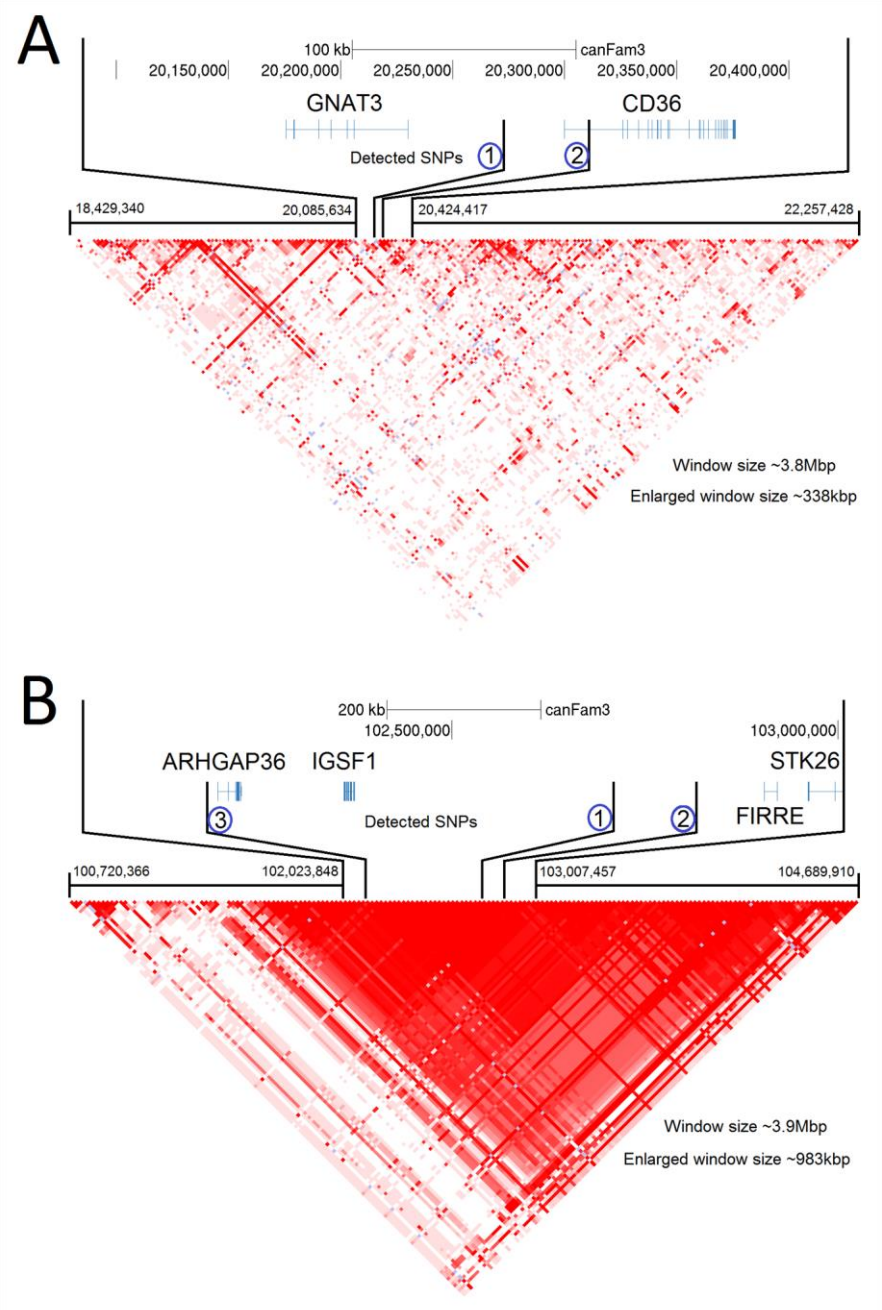

A
100 kb
canFam3
20,150,000
20,200,000
20,250,000
20,300,000
20,350,000
20,400,000
GNAT3
CD36
Detected SNPs
1
2
18,429,340
20,085,634
20,424,417
22,257,428
Window size ~3.8Mbp
Enlarged window size ~338kbp
B
200 kb
canFam3
102,500,000
103,000,000
ARHGAP36
IGSF1
STK26
FIRRE
3
Detected SNPs
1
2
100,720,366
102,023,848
103,007,457
104,689,910
Window size ~3.9Mbp
Enlarged window size ~983kbp

Fig. 6

	C-BARQ Trait	C-BARQ Trait Description[1] (number of questionnaire items)
Aggression	Stranger-directed	Severity of threatening or aggressive responses to strangers approaching or invading the dog's or owner's personal space, territory, or home range (10)
	Dog-directed	Severity of threatening or aggressive responses when approached directly by unfamiliar dogs (4)
	Owner-directed	Severity of threatening or aggressive responses to the owner or other members of the household when challenged, manhandled, stared at, stepped over, or when approached while in possession of food or objects (8)
	Dog Rivalry	Severity of aggressive or threatening responses to other familiar dogs in the household (4)
Fear	Stranger-oriented	Severity of fearful or wary responses when approached directly by strange or unfamiliar people (4)
	Dog-oriented	Severity of fearful or wary responses when approached directly by unfamiliar dogs (4)
	Nonsocial	Severity of fearful or wary responses to sudden or loud noises, traffic, and unfamiliar objects and situations (6)
	Separation-related behavior (anxiety)	Frequency of vocalizing and/or destructive behavior when separated from the owner, including autonomic signs of anxiety—restlessness, loss of appetite, trembling, and excessive salivation (8)
	Touch Sensitivity	Severity of fearful or wary responses to potentially painful or uncomfortable procedures, including bathing, grooming, nail-clipping, and veterinary examinations (4)

Table S1. C-BARQ fear and aggression trait descriptions (numbers of questionnaire items in parentheses)

Breed[1]	Trait	N	Mean	StDev	Var
BeT	DogAggr	110	0.908	0.814	0.663
BeT	DogFear	110	0.616	0.821	0.674
BeT	NonFear	112	0.616	0.646	0.417
BeT	OwnAggr	112	0.118	0.281	0.079
BeT	RivalryAg	107	0.635	0.632	0.399
BeT	SepFear	112	0.364	0.493	0.243
BeT	StrAggr	112	0.640	0.542	0.293
BeT	StrFear	111	0.632	0.889	0.791
BeT	TouchFear	111	0.501	0.480	0.231
BMD	DogAggr	195	0.506	0.676	0.456
BMD	DogFear	194	0.500	0.689	0.474
BMD	NonFear	195	0.622	0.678	0.460
BMD	OwnAggr	195	0.061	0.195	0.038
BMD	RivalryAg	185	0.250	0.544	0.296
BMD	SepFear	194	0.261	0.403	0.162
BMD	StrAggr	194	0.349	0.397	0.158
BMD	StrFear	196	0.619	0.905	0.818
BMD	TouchFear	193	0.392	0.469	0.220
BoC	DogAggr	640	1.068	0.990	0.980
BoC	DogFear	649	0.846	0.863	0.746
BoC	NonFear	650	0.965	0.792	0.627
BoC	OwnAggr	659	0.135	0.362	0.131
BoC	RivalryAg	588	0.667	0.810	0.656
BoC	SepFear	661	0.472	0.615	0.378
BoC	StrAggr	649	0.596	0.678	0.460
BoC	StrFear	647	0.731	0.940	0.884
BoC	TouchFear	643	0.750	0.790	0.624
BoT	DogAggr	70	1.289	1.125	1.265
BoT	DogFear	72	0.981	0.932	0.868
BoT	NonFear	70	0.716	0.677	0.459
BoT	OwnAggr	72	0.097	0.209	0.044
BoT	RivalryAg	60	0.281	0.633	0.401
BoT	SepFear	2	0.520	0.532	0.283
BoT	StrAggr	71	0.405	0.462	0.214
BoT	StrFear	72	0.524	0.845	0.713
BoT	TouchFear	71	0.720	0.686	0.470
BrS	DogAggr	111	0.673	0.778	0.606
BrS	DogFear	109	0.664	0.764	0.583
BrS	NonFear	112	0.572	0.498	0.248
BrS	OwnAggr	116	0.094	0.249	0.062
BrS	RivalryAg	97	0.353	0.509	0.259
BrS	SepFear	115	0.661	0.617	0.381
BrS	StrAggr	115	0.438	0.481	0.232

BrS	StrFear	116	0.460	0.672	0.451
BrS	TouchFear	116	0.646	0.673	0.453
Elk	DogAggr	33	0.725	0.887	0.787
Elk	DogFear	33	0.606	0.862	0.742
Elk	NonFear	33	0.631	0.644	0.414
Elk	OwnAggr	33	0.225	0.518	0.268
Elk	RivalryAg	27	0.241	0.329	0.108
Elk	SepFear	33	0.331	0.494	0.244
Elk	StrAggr	33	0.605	0.738	0.544
Elk	StrFear	33	0.318	0.567	0.321
Elk	TouchFear	32	0.690	0.695	0.483
ESt	DogAggr	76	0.602	0.827	0.684
ESt	DogFear	74	0.720	0.877	0.770
ESt	NonFear	76	0.490	0.562	0.316
ESt	OwnAggr	78	0.115	0.317	0.101
ESt	RivalryAg	67	0.578	0.722	0.521
ESt	SepFear	78	0.423	0.577	0.333
ESt	StrAggr	78	0.316	0.345	0.119
ESt	StrFear	78	0.327	0.538	0.289
ESt	TouchFear	75	0.421	0.588	0.346
Eur	DogAggr	45	0.461	0.538	0.290
Eur	DogFear	48	0.293	0.323	0.105
Eur	NonFear	48	0.453	0.420	0.176
Eur	OwnAggr	47	0.118	0.358	0.128
Eur	RivalryAg	44	0.470	0.718	0.515
Eur	SepFear	48	0.271	0.435	0.189
Eur	StrAggr	48	0.382	0.474	0.225
Eur	StrFear	45	0.517	0.728	0.530
Eur	TouchFear	47	0.576	0.568	0.322
FSp	DogAggr	10	1.067	0.700	0.489
FSp	DogFear	10	0.708	0.429	0.184
FSp	NonFear	10	0.920	0.658	0.433
FSp	OwnAggr	10	0.101	0.277	0.077
FSp	RivalryAg	7	1.214	0.847	0.717
FSp	SepFear	10	0.313	0.531	0.282
FSp	StrAggr	10	0.350	0.409	0.167
FSp	StrFear	10	0.500	0.874	0.764
FSp	TouchFear	10	0.700	0.550	0.303
GoS	DogAggr	27	0.784	0.784	0.614
GoS	DogFear	27	0.657	0.896	0.804
GoS	NonFear	27	0.618	0.556	0.309
GoS	OwnAggr	27	0.061	0.107	0.011
GoS	RivalryAg	26	0.635	0.686	0.471
GoS	SepFear	27	0.482	0.717	0.514

GoS	StrAggr	26	0.450	0.502	0.252
GoS	StrFear	27	0.444	0.864	0.747
GoS	TouchFear	27	0.521	0.528	0.278
Gry	DogAggr	138	0.536	0.798	0.636
Gry	DogFear	139	0.758	0.864	0.747
Gry	NonFear	141	1.007	0.849	0.721
Gry	OwnAggr	143	0.132	0.261	0.068
Gry	RivalryAg	120	0.526	0.628	0.394
Gry	SepFear	138	0.629	0.701	0.491
Gry	StrAggr	138	0.138	0.322	0.104
Gry	StrFear	139	0.698	1.077	1.159
Gry	TouchFear	139	0.690	0.759	0.577
IrW	DogAggr	45	0.541	0.974	0.948
IrW	DogFear	44	0.472	0.791	0.626
IrW	NonFear	45	0.660	0.679	0.461
IrW	OwnAggr	45	0.045	0.094	0.009
IrW	RivalryAg	38	0.423	0.758	0.574
IrW	SepFear	45	0.391	0.444	0.197
IrW	StrAggr	45	0.210	0.235	0.055
IrW	StrFear	45	0.411	0.745	0.554
IrW	TouchFear	44	0.511	0.604	0.365
JRT	DogAggr	284	1.477	1.142	1.304
JRT	DogFear	279	0.826	0.906	0.821
JRT	NonFear	285	0.930	0.764	0.584
JRT	OwnAggr	291	0.322	0.547	0.299
JRT	RivalryAg	242	0.983	0.958	0.918
JRT	SepFear	291	0.707	0.668	0.446
JRT	StrAggr	286	0.799	0.770	0.592
JRT	StrFear	289	0.575	0.837	0.701
JRT	TouchFear	279	0.833	0.715	0.511
NFd	DogAggr	45	0.674	0.782	0.611
NFd	DogFear	48	0.462	0.653	0.426
NFd	NonFear	49	0.712	0.643	0.413
NFd	OwnAggr	49	0.131	0.355	0.126
NFd	RivalryAg	42	0.448	0.603	0.364
NFd	SepFear	49	0.455	0.431	0.186
NFd	StrAggr	49	0.350	0.393	0.154
NFd	StrFear	49	0.199	0.635	0.404
NFd	TouchFear	49	0.617	0.611	0.373
NSD	DogAggr	66	0.607	0.716	0.512
NSD	DogFear	66	0.525	0.658	0.434
NSD	NonFear	66	0.701	0.672	0.452
NSD	OwnAggr	67	0.091	0.208	0.043
NSD	RivalryAg	62	0.496	0.612	0.374

NSD	SepFear	67	0.465	0.496	0.246
NSD	StrAggr	65	0.299	0.354	0.125
NSD	StrFear	66	0.451	0.603	0.364
NSD	TouchFear	67	0.657	0.681	0.464
Sci	DogAggr	27	1.016	0.921	0.848
Sci	DogFear	27	0.633	0.681	0.464
Sci	NonFear	27	0.595	0.549	0.301
Sci	OwnAggr	27	0.186	0.347	0.120
Sci	RivalryAg	26	0.667	0.759	0.577
Sci	SepFear	26	0.351	0.385	0.148
Sci	StrAggr	26	0.764	0.717	0.513
Sci	StrFear	26	0.538	0.673	0.453
Sci	TouchFear	26	0.788	0.848	0.719
ShP	DogAggr	37	1.392	1.008	1.016
ShP	DogFear	36	0.701	0.795	0.632
ShP	NonFear	38	0.735	0.760	0.577
ShP	OwnAggr	40	0.238	0.504	0.254
ShP	RivalryAg	26	0.772	0.753	0.567
ShP	SepFear	40	0.699	0.818	0.670
ShP	StrAggr	40	1.029	0.866	0.750
ShP	StrFear	38	0.961	0.907	0.823
ShP	TouchFear	39	1.077	1.081	1.169
Wei	DogAggr	107	0.805	0.754	0.568
Wei	DogFear	109	0.878	0.792	0.628
Wei	NonFear	107	0.612	0.669	0.448
Wei	OwnAggr	112	0.123	0.401	0.160
Wei	RivalryAg	97	0.490	0.654	0.427
Wei	SepFear	111	0.799	0.845	0.714
Wei	StrAggr	109	0.641	0.683	0.467
Wei	StrFear	111	0.585	0.776	0.602
Wei	TouchFear	109	0.616	0.687	0.473

Table S2. C-BARQ behavioral phenotypes used to test predictive model

[1]Breed abbreviations from Vaysse et al. (*1*)

References

1. A. Vaysse *et al.*, Identification of Genomic Regions Associated with Phenotypic Variation between Dog Breeds using Selection Mapping. *PLoS Genet* **7**, e1002316 (2011).

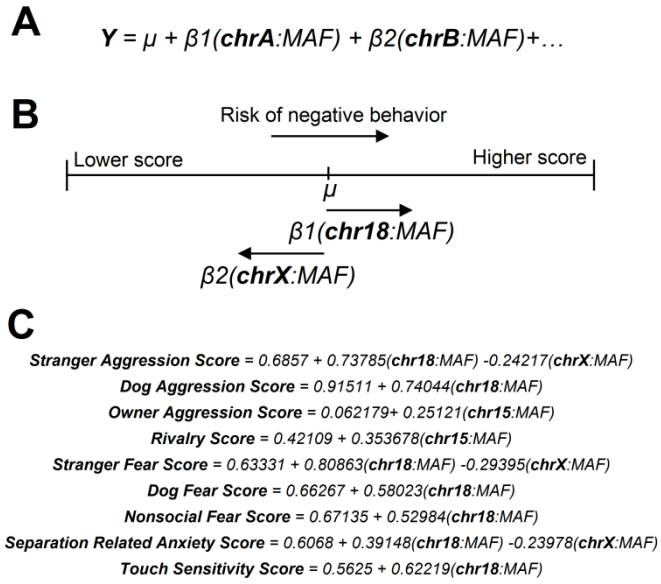


Fig. S1. Summary of prediction equation. (A) Generalized equation of the model used for the prediction of behavioral traits. MAF corresponds to the Major Allele Frequency of the SNP marker evaluated. (B) Interpretation example of the equation parameters for the prediction of negative behavior. μ sets the overall mean, β1(chr18:MAF) increases the risk while β2(chrX:MAF) decreases the risk. Therefore, a dog breed with high risk for negative behavior has a high chr18:MAF and a low chrX:MAF; on the other side, a dog breed with low risk for negative behavior has a low chr18:MAF and a high chrX:MAF. This specific example applies directly to Stranger-directed aggression, Stranger-oriented fear and Separation-related anxiety. (C) Regression equations for each of the nine C-BARQ behavioral traits.

Appendix 6

Analysis of novel gene-environment interactions possibilities: Draft manuscript: Zapata and Alvarez

See discussion of rationale for relevance to cancer risk and this project in Appendix 3.

Title: Fine-mapping and biological relevance implicate *GNAT3* and *IGSF1* in canine fear and aggression

Authors: Isain Zapata[1] and Carlos E. Alvarez[1,2,3]*

Affiliations:
[1]Center for Molecular and Human Genetics, The Research Institute at Nationwide Children's Hospital, Columbus, OH, 43205, USA
[2]Department of Pediatrics, The Ohio State University College of Medicine, Columbus, OH, 43210, USA
[3]Department of Veterinary Clinical Sciences, The Ohio State University College of Veterinary Medicine, Columbus, OH, 43210, USA
*Corresponding Author: alvarez.73@osu.edu

One Sentence Summary: Analysis of genome scans, positive selection and biological relevance implicate *GNAT3* and *IGSF1* in canine fear and aggression.

Abstract:
While extensive linkage disequilibrium (LD) in dogs facilitates broad mapping, it is a formidable obstacle for fine mapping. We combine data from interbreed genome scans and intrabreed signatures of positive selection to fine-map dog traits. After demonstrating the concept with two known canine-trait loci, we analyzed two new genome loci associated with canine fear and aggression. We thus implicated *GNAT3* (chr18) and *IGSF1* (chrX) to be the top genes associated with fear and aggression in dogs. Our ability to fine map the X chromosome variant is striking because it lies within a ~3Mb LD block that has the highest signal of selection in dogs and also contains association peaks for increased size (which is in perfect LD with the decreased-fear/aggression allele in 9/11 tested breeds) and sociability (which we suggest is a distinct trait associated with *HS6ST2*). The patterns of brain expression of *GNAT3* and *IGSF1* indicate they are excellent candidates for involvement in the fear pathway centered on the amygdala/ hypothalamic-pituitary-adrenal axis.

Main Text:

Introduction

The dog is an extremely powerful genetic model in unique ways (*1, 2*). Despite earlier divergence time with humans than mouse, dogs are more similar to humans at the levels of DNA and protein sequence – and share 600Mb of DNA sequence that was lost in mice (*1, 3*). Dogs age 5-7 times faster than humans, are numerous (~75M in the US), receive health care and have several hundreds of characterized inherited diseases. But most importantly, dogs represent a one-of-a kind genetics resource. There are approximately

400 dog breeds and each is on the order of 100-fold simpler than the full population. Within breeds, variation is far less than within the major human ethnic groups; but across breeds, variation is much greater than across human populations. Dogs are thus powerful for broad mapping in genome scans, but have two potential limitations: i) fine mapping is difficult due to expansive linkage disequilibrium, or LD (on the order of 50-fold greater than in humans (*4*)); and ii) some traits cannot be mapped by classical genetic approaches within single breeds because each breed is fully or nearly homozygous over approximately 25% of the genome (*5*). Successful studies have mitigated those problems by first broad mapping across breeds and, after discovery of additional variation in the region of interest, fine-mapping within multiple related breeds that are known to share a founder variant.(*6*) This approach has in part given way to the use of increased marker densities – currently ~175,000 single nucleotide polymorphism (SNP) markers – to enable finer mapping in a single study. But the high levels of LD in dogs continue to be a major limitation to fine mapping. In such cases, experimental screening of candidates may not be feasible and the only recourse may be prioritization based on variant analysis and biological relevance (*7*).

New studies indicate that there was a single origin of dog domestication from wolves approximately 33,000YA in southern East Asia; and this was followed by migration to the Middle East, Africa and Europe, beginning about 15,000YA (*8*). The evolutionary history of dogs has led several investigators to search for evidence of selection across the genome (as has been done in humans (*9*)) (*5*, *10-13*). For instance, Akey et al. identified a very large region of population differentiation in the Shar-Pei and showed that it is associated with its unique trait of wrinkled skin. Similar patterns of selection have also been reported in cases where a breed stereotype is associated with a fully penetrant haplotype that is often fixed for the risk or non-risk allele within any breed (*5*, *12*). The landmark work by Boyko et al. and Vaysse et al. represent the most comprehensive studies of stereotype mapping and population genetics across breeds. The former included 915 dogs from 80 dog breeds and the latter 509 dogs from 46 breeds; there was partial breed overlap in the two (these genotype resources are the basis for the present study). Vaysse et al.'s article is titled "Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping" (*5*). In addition to genomewide association (GWA) mapping of boldness, sociability and several morphological breed-stereotypes, they conducted extensive discovery of evolutionary footprints of selection. That included population differentiation using the fixation index (F_{ST}; when several breeds share founder variation) and D_i (designed to detect selection in one or few breeds out of a larger group; pairwise F_{ST} values are normalized for a breed vs. the genomewide average, then summed across pairwise combinations involving that breed (*10*)), extended haplotype homozygosity (EHH, or cross-population EHH; the probability of identity by descent is calculated using all genotypes in a dataset (*9*)), and reduced heterozygosity (S_i; the sum of regional deviations in levels of genomewide relative-heterozygosity between two breeds is compared to the genomewide average, and the sum of those across all pairwise comparisons is calculated (*5*)). Of Vaysse et al.'s top 18 F_{ST} regions, 10 overlapped a total of 15 GWA loci that were previously published (n=11) or new (n=4). The top two F_{ST} regions are approximately 2.5 and 2Mb long, and

each contains at least three known trait associations that are distinct according to haplotype analysis. Boyko et al. had similar result, with seven of the top 10 F_{ST} regions being the same in both studies. In addition, Vaysse et al. identified extensive evidence of selective sweeps that may be specific to one or a few breeds, which could be missed by F_{ST} analysis. Several regions of high differentiation (D_i) and reduced heterozygosity (S_i) correspond to known stereotype-associated haplotypes (e.g., skin wrinkling, chondrodysplasia and furnishings), but the vast majority of those regions have not been linked to any trait.

Genetic hitchhiking generally refers to neutral variation that is carried along with an allele under positive selection (*14*). As the selected variant goes to fixation, there is a loss of variation flanking that site (termed a selective sweep) (*15, 16*). LD in such a selected region increases dramatically, and that effect is the basis of several approaches for identifying positive selection (e.g., EHH, which detects large haplotypes suggestive of selective sweeps). In contrast to the early LD effects, as the functional allele approaches a frequency of 50%, a distinct pattern emerges: LD on one side of the selected allele will remain high and decrease slowly, but LD across the selected site decreases quickly (*17, 18*). Because dog breeds represent hundreds of isolated populations, they are an ideal species in which to use hitchhiking patterns of LD to fine-map variation that underwent selection. Some breeds should have strong LD only upstream of the selected site, and others only downstream. Although hitchhiking theory is rarely mentioned there, its signature of extended LD can be seen in several reports where one or few breeds exhibit a very large block of homozygosity. Most of those loci were first mapped by GWA to some trait (e.g., chondrodysplasia and furnishings), with the exception of wrinkled skin in the Shar-Pei (*10, 19, 20*). Here we directly test for evidence for and against selection at canine peak GWA-SNPs using the conventional approaches of phasing or identity by descent (IBD) haplotype analysis. These have been used in human genetics in the contexts of population genetics and IBD mapping (*21*). Until the advent of high throughput sequencing, limited human haplotype-phasing was implemented through triad analysis or direct experimental methods. Rather, most of that work was and continues to be done through probabilistic methods.

We show that haplotype phasing can be done within breeds by separately analyzing all dogs homozygous for each allele at any position (and determining the extent of flanking homozygosity). We first demonstrate the use of D_i, S_i and haplotype phasing analysis applied to evaluate previously published GWA loci, and then use it to fine map variation identified by GWA of canine fear and aggression traits in two cohorts with partially overlapping breeds (*22*). Specifically, single loci on chr18 and chrX are associated with stranger-directed fear, dog-directed fear, stranger-directed aggression, dog-directed aggression, touch-sensitivity and non-social fear; and the chrX locus is also associatated with separation anxiety. Of those traits, dog-directed fear, touch sensitivity, non-social fear and separation anxiety are also associated with known variants at *IGF1*, *IGF1R* or *HMGA2* that are associated with small size. However, in contrast to the chr18 and X variants, those associations had a negligible effect in predictive models that were validated in a third group of dogs from different breeds. Rather, small size variation is strongly predictive for owner-directed aggression and dog rivalry (which are not strongly

associated with the chr18 and X loci). By combining that chr18 and X GWA data with multiple types of evidence of selective sweeps and the gene expression patterns of a total of five genes at those loci, we implicate two genes in the core neuroanatomical circuit that drives fear and aggression – the Hypothalamic-Pituitary-Adrenal (HPA) Axis.

Results

Signatures of positive selection

The present studies were devised to fine map two loci from our GWA study of canine fear and aggression traits (*22*). Our mapping was done using the published breed behavioral stereotypes from the C-BARQ resource and the genotype datasets of Boyko et al. and Vaysse et al. (*5, 12, 23*). The latter studies have partial overlap in breed makeup; Boyko's has more breeds and dogs (genotyped on the Affymetrix v.2 Canine array), and Vaysse's has higher SNP density (Illumina CanineHD array). Both of our primary fear and aggression loci lie among the highest F_{ST} regions in dogs. The top F_{ST} region in both studies, on chrX, includes one of our GWA peaks. That interval was said to be 3.7Mb (0.795 F_{ST}) and 2.6Mb (0.75 F_{ST}) long, respectively (*5, 12*). Boyko et al. reported this region is associated with body size and many aspects of skull shape (both absolute and proportional; distinct from brachycephaly), but these were only reported as r^2 correlation values for the general chrX:104Mb region (without mention of peak SNPs or p-values). In the present work, we confirmed this chrX association with size, but using the Vaysse genotype dataset (see Methods; (*5*)), and show the location of the peak SNP. Vaysse et al. also mapped sociability within this chrX F_{ST} interval. The region is thus consistent with the increased LD that is associated with a selective sweep, but interpretation and fine mapping are complicated by the reduced recombination rate on chrX. The other fear and aggression locus lies on the chr18 F_{ST} region that ranks fifth in Boyko's analysis, but is absent using the same thresholds in Vaysse (F_{ST} >0.55; minor allele frequency >15%). This single-SNP F_{ST} site tags, by linkage, the *FGF4* retrogene insertion that causes chondrodysplasia in several breeds (but some breeds have that SNP allele/ancestral haplotype without the insertion). Because it is absent in Vaysse's F_{ST} regions (which has higher SNP density), this indicates that our fear/aggression GWA hit in both datasets is unlikely to be a false positive due to shared population structure in the two cohorts. We found no report of attempts to use genetics or biology to identify the functional variation on the chrX locus. This seems surprising given that it is the top F_{ST} locus across different samples of diverse dog breeds and that it is the first known locus associated with large size and an important behavioral trait (sociability). By first appearances the region seems impenetrable by fine-mapping due to strong and extensive LD. The sociability trait (which was only significant in Vaysse's analysis of female dogs) maps to a peak of 10 SNPs spanning 580kb.

We first conducted analysis of known GWA haplotypes for signatures of ancient and recent selection. The small-size functional variant in *IGF1*, which has the strongest contribution to dog size across all but very large dog breeds, was indicated to predate dog domestication (*10, 11*). Figure 1A shows contiguous SNPs in the Vaysse dataset centered on the GWA peak SNP for small size that lies within *IGF1* on chr15 (*5*). Thirteen of the 28 breeds have D_i regions that intersect the peak SNP, and those represent both small and

large breeds. The strong differentiation signal is consistent with Boyko and Vaysse detecting this locus as the fourth and fifth top F_{ST} region, respectively. S_i regions from two breeds intersect the tag marker, and those from five others comprise a small block upstream. We interpret this observation to mean that the reduced variation associated with the selective sweep is negligible because the selection went to fixation so long ago. The question remains why seven non-small breeds share a block of S_i shortly upstream of the small allele (e.g., this could reflect selection of another functional variant at this locus). We also phased haplotypes in individual breeds by selecting all homozygotes of each allele at the peak SNP and then separately measuring the length of homozygosity across all individuals. Put another way, we directly determined phasing rather than using probabilistic methods. The pattern of phased haplotypes suggests the T allele haplotype that underwent selection for small size long ago is still generally longer than the alternative allele's haplotype (18 and 15 kb, respectively). The Rottweiler was previously shown to be anomalous because it has a high frequency the small dog allele, but is in fact very large (*24*); to us this is consistent with a second variant strongly compensating for the small dog variant. All three tests indicate signal in the Rottweiler, including the longest S_i region and phased haplotype of all the breeds. Curiously, all eight breeds with large phased haplotypes overlap the upstream seven-breed block of S_i. All of those S_i breeds for which we have phasing (5 of the 7) have large phased haplotypes, and that includes the four largest – which range from about 200-400kb.

Next we looked at the furnishings variant within *RSPO2* on chr13, which was selected in different breeds relatively recently (Fig. 1B) (*20*). Five breeds carry the peak GWA SNP associated with furnishings, all of which are known to have the phenotype (furnishings on Jack Russell Terrier is evident in the long coated version recognized by the US American Kennel Club – called Parson Russell Terrier). Of those five breeds, four have D_i signal spanning the peak SNP (a fifth is shifted ~50kb upstream) and two of those same four have S_i signal (two others are shifted ~20 and 70kb). [Only one non-furnishings breed has overlapping D_i signal and it is the smallest D_i interval of all, and none has S_i signal.] Phasing analysis shows the haplotype containing the selected allele is much larger than the alternative allele's (range of 186-1,369 vs. 7.5-43kb; minimal overlapping intervals of 186 vs. 7.5kb). To contrast these findings for anciently and recently selected alleles, we evaluated polymorphic alleles that are not well implicated by signatures of selection (Supplementary Fig. 1). These show significant differences in patterns and range of sizes of phased haplotypes. Thus, future studies are necessary to consider the feasibility and utility of genome scanning for measures of positive selection including phased haplotype analysis.

In Figure 2 we present analysis of the new fear and aggression locus on chr18 (*22*). The haplotype phasing analysis shows the two breeds with the increased risk allele – Dachshund and Yorkshire Terrier – have much larger haplotypes compared to the alternative allele (ranging up to 683 vs. 186; minimal overlap regions of 418 vs. 13.2 kb). Their haplotype sizes, and the central position of the putative variant under selection, suggest a recent selective sweep according to hitchhiking theory. Notably, only a subset of the increased-fear/aggression haplotype under selection contains the *FGF4* retrogene insertion that causes chondrodysplasia in some breeds but not in others. All known

exceptions to the co-occurrence of the retrogene insertion and the phenotype are very small dogs – Yorkshire Terrier, Chihuahua and Japanese Chin – leading the authors to propose the trait is not manifest in carriers that are small (*19*). [That study included four Yorkshire terriers that had an insertion-allele frequency of 50%; in our phasing analysis of this breed, we found 4 dogs homozygous for the increased-fear allele (Fig. 2) and 3 homozygous for the alternative allele (which reveals a short phased haplotype spanning 8 SNPs).] Moreover, the existence of insertion-positive and -negative versions of the ancestral chr18 increased-fear/aggression haplotype are also evident from Cocker Spaniels and Beagles – both of which have >20% allele frequencies of the increased-fear/aggression haplotype (*5*), but lack the insertion and chondrodysplasia (*19*). Shar-Pei's and Huskies also have >20% frequencies of the increased-fear/aggression haplotype (*5, 12*), but have negligible or no chondrodysplasia. Thus, there is evidence of recent selection for the increased-fear/aggression haplotype in Dachshunds and Yorkshire Terriers (large haplotype size); that selection is presumed to be for short legs in the former but unknown for the latter. Only the two breeds that carry the chr18 increased-fear/aggression allele at high frequency have D_i signal spanning the peak SNP, and both also have the longest S_i signal either intersecting or close to the peak SNP. The allele protective of fear and aggression has high allele frequencies in all but the two breeds that carry the risk allele, and is fixed (>95%) in 22 of the 30 breeds (5 others are ≥75%). In contrast, two sources of wolf sequence data from around the world show the wolf allele frequency of the two alleles to be ~50%. We propose these patterns are consistent with selection during the domestication process. If so, that selected reduced-fear and aggression variant is narrowed to a two-SNP minimal overlap interval of phased haplotypes.

In Figure 3 we present a 5Mb chrX region that contains GWA peaks for size, fear and aggression, and sociability. As described above, this region has the highest F_{ST} level in the dog genome and has strong LD across 2.6-3.7Mb (*5, 12*). The overall pattern of phased haplotypes and D_i and S_i signal suggests the three traits may be distinct. The fear/aggression peak overlaps D_i and S_i signal for many breeds, but that for size and sociability do not. However, there is perfect LD between the size and fear/aggression peaks in this haplotype in 9 of 11 breeds, and more data is necessary to establish their relationship. Human mutations in the top chrX candidate gene for fear/aggression, *IGSF1*, are known to affect human Growth Hormone biology (*25*), suggesting the two dog traits could share all or some genetic variation at this locus; additionally, expression of the gene abutting the peak SNP for size, *ARGHAP36*, is strongly correlated with *IGSF1* expression (see below). Although all but one of the breeds with the fear/aggression-protective allele have perfect LD with the sociability allele, only half of the breeds with the sociability allele have the fear/aggression-protective allele in perfect LD. All breeds with the reduced-fear/aggression allele show D_i signal overlapping the fear/aggression peak, but there is only one segment of each overlapping the sociability peak. On the other hand, there are eight S_i segments at the fear and aggression peak (one is shifted a single SNP away), but none is in a breed with the fear/aggression-protective allele. Thus, the predominant signatures of positive selection point to the association with fear and aggression. This is also suggested by the phased haplotype analysis, which

shows minimal overlapping regions of 841 vs. 284 kb for the fear/aggression-protective and alternative alleles, respectively. In contrast, the appearance of such an effect at the sociability locus is due to the contribution of the 10 breeds that carry the fear/aggression-protective allele in perfect LD; and the minimal overlapping regions there are 85 vs. 47 kb for the social allele and the non-social allele, respectively. Notably, this locus presents the possibility of a special case of hitchhiking in which selection initially occurs on one functional variant (here reduced-fear/aggression), but a second variant in the region is also favored in subsequent selection (here, size or sociability). As a result, relative to the original variant selected, the LD breaks down only on the side opposite the second selected variant.

Based on modern worldwide wolf sequences, the minor chrX allele is that associated with increased-fear/aggression (43% frequency) (*26*). Of the diverse 30 breeds studied here, 20 have frequencies >70% for the increased-fear/aggression allele and 10 have such frequencies for the alternative allele (with many breeds fixed for one of those). In Discussion we present the considerations for interpreting this.

Fine-mapping functional variation

Based on our test cases of known trait-associated variants selected in ancient or recent times, we called the minimal intervals for fear and aggression on chr18 and X. Vaysse et al. showed that top ranked S_i and D_i signals mirror each other in a subset of regions, but the effects are highly variable; and interpretation seems complex and lacking fine resolution. The furnishings locus on chr13 is an example of complementing S_i and D_i information (Fig. 1B). However, whereas all five breeds with furnishings have D_i intervals that intersect the peak SNP, three of those lack such evidence for S_i. D_i may thus be more sensitive and precise for fine mapping in general, but many regions appear to be more broadly flagged by both statistics. Based on the findings above, we propose the following scheme: i) the maximum map interval corresponds to the region of overlap of phased haplotypes across all breeds (separately applied to both SNP alleles at a GWA peak); and ii) the minimum map interval can combine that with the smallest region of overlap of D_i or S_i regions across breeds. Using these criteria, the ancient small-size variant in *IGF1* is correctly predicted to lie in a maximum map interval of 18 kb, which is the same as the minimum map interval since D_i or S_i regions do not further delimit the region. For the recently selected furnishings-variant in *RSPO2*, it is correctly predicted to lie in a maximum map interval of 186 kb and a minimum map interval of 59 kb. In accordance with hitchhiking theory, it seems likely that the maximum map interval for a region under selection will always capture the functional variant. However, it is not clear how precisely that critical region could be narrowed using different signatures of selection; each case could be unique and is likely to require additional confirmation.

For the chr18 fear and aggression locus, the predicted maximum map interval for the most recently selected increased-fear/aggression haplotype is 418 kb and the minimum is 108 kb. However, if there was initially selection for the reduced-fear/aggression allele at domestication and subsequent selection for the increased-fear/aggression allele at the same position (as we speculate based on parsimony), then the maximum and minimum map intervals would be 13.2 kb. For the chrX locus, the

predicted maximum map interval for the reduced-fear/aggression allele is 841 kb, and the minimum map interval is 218 kb. The sociability locus lacks significant D_i/S_i evidence and has no phenotype for a key breed. The selected sociability allele T thus has maximum and minimum map intervals of 147 kb. Notably the original report of this sociability GWA, using the same genotype data, was only suggestive in the full dataset; rather, genomewide significance was detected in analysis of females exclusively, and yielded a peak of 10 SNPs spanning 580 kb.

Biological relevance of candidate genes

According to the minimal intervals, we have implicated the following genes: i) in the chr18 fear and aggression locus, *GNAT3* and *CD36*; ii) in the chrX fear and aggression locus, *IGSF1*, *FIRRE* (long non coding RNA) and STK26; and iii) in the chrX sociability locus, *MBNL3* and *HS6ST2*. Both sociability candidates are expressed in the brain: *MBNL3* at very low levels and apparently not at all enriched in brain, but *HS6ST2* at very high levels and highly enriched in many brain regions (BioGPS microarray data for 176 human and 191 mouse tissues and cells (*27*); NCBI GEO GSE1133 and GSE10246). We favor *HS6ST2* for the sociability association. For the fear and aggression loci, we only see clear biological relevance (mainly expression in sensory organs, brain or adrenal gland) for *GNAT3* and *IGSF1*.

GNAT3 encodes Gustducin alpha, the G alpha subunit that transduces taste receptor signaling. Gustducin alpha also has chemosensory roles in the vomeronasal organ, airways and gastrointestinal tract (*28, 29*). There are also reports that *GNAT3* is expressed in areas of the brain that include the brainstem, hypothalamus and hippocampus (*30*-*32*). However, those studies targeted specific brain regions and we are not aware of systematic analysis of the entire brain in any mammal. We consulted the Allen Brain Atlas of *in situ* mRNA hybridization analysis (*33*) and found evidence that *GNAT3* is most highly expressed in the amygdala in the adult mouse, specifically in layer 2 of the Cortical Amigdalar Area ("CoA"; Fig. S2). That finding is supported by analysis of public gene expression data showing that the highest ranked expression-change of *Gnat3* mRNA in any brain region is a 3.73-fold increase in the amygdala of rats 6h after pain exposure (Nextbio analysis server (*34*): *Gnat3* has a percentile gene-ranking score=99 for the microarray experiment, p=0.0165; NCBI GEO Accession GSE1779 (*35*)). The next-highest ranked (and significant after multiple testing correction) gene expression-changes reported in experimental paradigms involving the brain follow: i) rat hippocampus (GSE3531, score=82); ii) mouse cortex (GSE31840, score=81); and iii) mouse striatum (GSE48955, score=65). In addition to very high expression in the amygdala, the Allen Brain Atlas of the adult mouse reveals lower levels of expression in parts of the pons: Lateral reticular nucleus, Paragigantocellular reticular nucleus, lateral part, and the Facial motor nucleus.

IGSF1 is Immunoglobulin superfamily member 1. Common human variation in *IGSF1* is associated with age at menarche, and loss-of function mutations cause a human syndrome of congenital hypothyroidism, macroorchidism, Prolactin and Growth Hormone deficiency, delayed pubertal testosterone and obesity (*25, 36*). It is expressed very highly in the anterior pituitary and hypothalamus, but also in the choroid plexus,

adrenal gland, pancreas, heart and skeletal muscle, fetal liver and testis (*37*). To generate quantitative data, we used the BioGPS GeneAtlas of genomewide gene expression in 191 mouse tissues and cells. The increased-expression levels, relative to the median for all tissues are as follows: pituitary, 1,060-fold; hypothalamus, 340; amygdala, 26; heart, 22; spinal cord, 11; hippocampus, 10; placenta, 10; and nucleus accumbens, 5.5 (data is for probe 1433652_at). Immunohistochemical evidence clearly shows that IGSF1 protein expression is restricted to neurons (including at low levels in the cortex, lateral ventricle, and cerebellum (*38*)). Notably, there is strong evidence that *IGSF1* is co-expressed with its two flanking genes, *ARHGAP36* and the long non-coding RNA *FIRRE*. In the mouse BioGPS GeneAtlas of 191-tissues/cells, expression of *IGSF1* (probe 1433652_at) has Pearson correlations of 0.85 with *FIRRE* (1436638_at; two other *FIRRE* probes also have correlations >0.76) and 0.80 with *ARHGAP36* (1454660_at); the same probe has a correlation 0.92 with a second *IGSF1* probe. The human BioGPS GeneAtlas of 176 tissues/cells lacks a probe for *FIRRE*, but *IGSF1* and *ARHGAP36* have a correlation of 0.97 (probes 207695_s_at and gnf1h04904_at). It is thus possible that variation in the *IGSF1* region could also affect the other two genes.

Discussion

Here we demonstrated the utility of combining widely used measures of positive selection to fine map GWA loci. This included breed differentiation, reduced variation and phased haplotype analysis. The latter was done similarly to how this is done in human triads: using a homozygous parent or offspring at any locus to determine the phased haplotype for it and, thus, for the heterozygous carrier. Here we analyzed each breed separately, and individually determined the extent of homozygosity spanning the risk and non-risk GWA alleles in homozygotes of each. One recurring pattern we see is that the selected haplotype tends to be larger, as expected. Another observed pattern that was expected, within a population (breed), is that the original LD that exists across the selected allele in a sweep breaks down over time; but LD will break down slowly between the selected allele and variants on one side of the selected site. Assuming the variation flanking a selected allele is neutral, then LD blocks (or phased haplotypes) from individual breeds will overlap at the middle but extend to one side or the other.

The phased-haplotype and D_i data for small size at the *IGF1* locus illustrate signatures of ancient selection. We interpret the chr18 fear/aggression locus to have similar phased-haplotype evidence of very old selection for the reduced-fear/aggression A allele. In addition there is recent selection of the increased-fear/aggression C allele in two breeds. Consistent with hitchhiking theory, the latter is indicated by the large phased-haplotypes in two breeds, and the fact that the putative variant under selection is in a central position within those haplotypes. In the Daschound, which is fixed for that allele in perfect LD with the chondrodysplasia *FGF4* retrogene insertion, that selection has been presumed to be for short legs. Yorkshire Terriers also have that insertion haplotype at high frequency but do not manifest proportionally-short legs – ruling out recent selection on that trait. [Other breeds, such as Beagles and Cocker Spaniels have high frequencies of the increased-fear allele, but lack the retrogene insertion and shortened legs.] While the Yorkshire Terrier was created in the 1800's for vermin control in mills

(and was also used for rat baiting), it is currently a very popular "toy" breed. That raises the possibility of recent human selection for dogs that prefer to be held closely (i.e., here, because they are fearful). Alternatively, both of these breeds could be under selection for aggression traits at the chr18 locus. In contrast to chr18, the chrX fear and aggression locus may be a special case in multiple respects – including that there may be a second linked allele under selection (for sociability or size), and that it is in a region with a low recombination rate (at least in part because it is on chrX). We also suspect that the increased- and decreased-fear/aggression variants at this same locus may both have been under selection at different times and in different breeds.

Most breeds have very high frequencies or are fixed for the chr18 reduced-fear/aggression allele. The phasing analysis of that haplotype is similar to that of the selected *IGF1* haplotype in small breeds: they are short and thus presumed to be ancient. These observations converge to implicate involvement of the chr18 decreased-fear/aggression allele in domestication. This is consistent with the widely held theory that dogs were domesticated by virtue of losing fear of humans (*39, 40*). In contrast to chr18, 20 of the diverse breeds studied here have the increased-fear/aggression chrX allele at >70% frequency, and only 10 have the reduced-fear allele at that level. Wolves have 57% allele frequency for the reduced-fear allele. The ancestral haplotype under selection is clearly that for decreased fear/aggression, but its age is difficult to discern because of its location on the X chromosome. Thus, despite the pattern of allele frequencies in the extant dog breeds and wolves sampled here, the putative advantage of reducing fear and aggression in the domestication of a predator leads us to speculate that selection of that allele was involved in domestication. The uneven distribution of the allele is consistent with geographical effects, gene flow between dogs and wolves, and natural or human selection. The latter possibility is hinted at by the few breeds that have long phased-haplotypes for the increased-fear/aggression allele, mainly the Weimaraner and Greyhound (both were originally hunting breeds). Also, it would not be surprising if wolf populations selected for the reduced-fear allele in the last several hundred years as interactions with humans became increasingly common. One possible scenario is that the ancestral wolf had a high frequency of the increased-fear/aggression allele; the alternative, reduced-fear, allele was selected in dog domestication (at least in some geographical areas); selection post-domestication favored the reduced fear allele in wolves; and natural/human selection or genetic drift resulted in the present patterns of the two alleles. The imminent availability of archeological/ancestral dog and wolf sequences should resolve this question (*41*).

Combining GWA mapping data with different signatures of selection, we fine-mapped two new loci for fear and aggression (*22*) and one previously reported for sociability using the same dataset (*5*). We reduced each locus to two coding genes (one locus also has a lncRNA). For sociability, we favor Heparan sulfate 6-O-sulfotransferase 2 (*HS6ST2*) because it is both very highly expressed and enriched in brain (the other candidate, Muscleblind-like 3 or *MBNL3*, is very weakly expressed and not enriched in brain). Biological relevance is further supported by reports that mice mutant for Exp1 heparan sulfate-synthesizing enzyme or Spock3 heparan sulfate proteoglycan both have impairments in social interactions (*42, 43*). The top candidate for fear and aggression on

chr18 is Gustducin alpha (*GNAT3*) which is a G alpha subunit that transduces signaling in taste cells and other chemosensory systems (including the vomeronasal organ), as well as in the brain. Because the fear and aggression traits include non-social fear (e.g., fear of loud sounds or novel objects), our working hypothesis is focused more on brain vs. chemosensory biology – particularly in the amygdala, hypothalamus, pituitary and hippocampus. The top candidate for fear on chrX is the gene Immunoglobulin superfamily member 1 (*IGSF1*) which causes a syndrome that includes congenital hypothyroidism and Growth Hormone deficiency when mutated in humans. The latter is interesting given that large dog size also maps to this same LD block (524 kbp downstream from our peak SNP for fear). While the biochemical function of IGSF1 is not well understood, it has multiple endocrine roles and is predominantly expressed in the pituitary and hypothalamus (also relevant here, it is expressed at much lower levels in the amygdala and adrenal gland). Studies of fear-relevant rat models showed that *GNAT3* has among the highest gene expression changes in the amygdala after a pain stimulus (described in Results; *35*), as *IGSF1* does in the cortex after both stress and tactile stimulation (*44*). In summary, we propose that regulatory variants in the neuronal genes *GNAT3* and *IGSF1* (and possibly the latter's flanking genes *ARHGAP36* and *FIRRE*, as the three genes are co-regulated) affect canine behavior by modulating the function of the central fear circuit that transduces sensory information through the amygdala to the hypothalamic-pituitary-adrenal axis.

These findings will require extensive experimental validation at the levels of canine genetics, neuroanatomy, physiology and behavior. Because the biochemistry and anatomy of fear and aggression are so highly conserved, it will also be possible to study these circuits in genetically-modified mice; but one must be mindful of the fact that mice are prey and dogs are predators. While the present study levearges the strengths of cross-breed genetics, it is now important to confirm and more precisely define the behavioral effects in breeds carrying both alleles at chr18 or X. This will lead to a meaningful understanding of these traits – including their development (*45*) and our ability to affect them behaviorally or pharmacologically (*46, 47*). Given the sparcity of knowledge of human fear/aggression genetics, these dog findings present rich opportunities for translational psychology and psychiatry (*48*).

Materials and Methods

Genotype, phenotype and GWA data

For the genotype data, we used a previously published SNP dataset (*5*). The dataset contains a total of 456 dogs from 30 breeds genotyped on the ~175,000 SNP Illumina CanineHD array. GWAS markers were used to define the traits based on breed stereotypes for size and furnishings (*5*), osteosarcoma (*49*), and fear and aggression (*22, 23*). Boyko et al. reported that a size variant lies somewhere in the 3.7Mb chrX region with the top-F_{ST} signal in the dog genome (*12*). We mapped the peak SNP in this region to chrX:105,245,495 using the the Vaysse genotypes (*5*) and the same GWA methods and significance thresholds as in our companion study (Genome-wide Efficient Mixed Model Association algorithm or GEMMA (*22, 50*)).

Phased haplotype analysis

We used the top GWA marker for each trait (or for arbitrarily chosen control regions) to segregate individuals within each breed by their carrier status: heterozygous and homozygous. Continuing the analysis within breeds, the homozygotes for alleles A and B were analyzed further. To construct the largest common phased haplotype within a breed, we defined the boundaries by evaluating each SNP upstream and downstream, and keeping only SNP markers that have an allele frequency of at least 0.95. The largest such interval for each allele at the peak SNP was called a phased-haplotype block. In this work we only include within-breed phased-haplotype blocks if they were present in four or more dogs. Allele frequency calculations and data analyses were performed in PLINK v.1.07 (*51*).

S_i/D_i blocks and gene annotation

We used S_i and D_i data previously published by Vaysse et al. using the same genotype data (*5*). This information is available as supplementary material at http://dogs.genouest.org/SWEEP.dir/Supplemental.html. We converted all SNP positions from the original CanFam2 assembly to CanFam3.1 to take advantage of the Broad Institute's CanFam3 Improved Annotation Data V.1 (*52*). The improved gene annotation is available as a public track on the UCSC genome browser.

References

1. C. E. Alvarez, Naturally Occurring Cancers in Dogs: Insights for Translational Genetics and Medicine. *ILAR Journal* **55**, 16-45 (2014).
2. J. J. Schoenebeck, E. A. Ostrander, Insights into Morphology and Disease from the Dog Genome Project. *Annual Review of Cell and Developmental Biology* **30**, 535-560 (2014).
3. K. Lindblad-Toh *et al.*, Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803-819 (2005).
4. N. B. Sutter *et al.*, Extensive and breed-specific linkage disequilibrium in Canis familiaris. *Genome Res.* **14**, 2388-2396 (2004).
5. A. Vaysse *et al.*, Identification of Genomic Regions Associated with Phenotypic Variation between Dog Breeds using Selection Mapping. *PLoS Genet* **7**, e1002316 (2011).
6. E. K. Karlsson *et al.*, Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat Genet* **39**, 1321-1328 (2007).
7. K. Kyöstilä, A. K. Lappalainen, H. Lohi, Canine Chondrodysplasia Caused by a Truncating Mutation in Collagen-Binding Integrin Alpha Subunit 10. *PLoS ONE* **8**, e75621 (2013).
8. G. D. Wang *et al.*, Out of southern East Asia: the natural history of domestic dogs across the world. *Cell Res*, (2015).
9. P. C. Sabeti *et al.*, Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-918 (2007).
10. J. M. Akey *et al.*, Tracking footprints of artificial selection in the dog genome. *Proceedings of the National Academy of Sciences* **107**, 1160-1165 (2010).
11. B. M. vonHoldt *et al.*, Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464**, 898-902 (2010).
12. A. R. Boyko *et al.*, A Simple Genetic Architecture Underlies Morphological Variation in Dogs. *PLoS Biol* **8**, e1000451 (2010).
13. F. Schlamp *et al.*, Evaluating the performance of selection scans to detect selective sweeps in domestic dogs. *Molecular ecology*, (2015).
14. J. M. Smith, J. Haigh, THE HITCHHIKING EFFECT OF A FAVORABLE GENE. *Genetical Research* **23**, 23-35 (1974).
15. N. L. Kaplan, R. R. Hudson, C. H. Langley, The "hitchhiking effect" revisited. *Genetics* **123**, 887-899 (1989).
16. W. Stephan, T. H. E. Wiehe, M. W. Lenz, The effect of strongly selected substitutions on neutral polymorphism: Analytical results based on diffusion theory. *Theoretical Population Biology* **41**, 237-254 (1992).
17. Y. Kim, R. Nielsen, Linkage disequilibrium as a signature of selective sweeps. *Genetics* **167**, 1513-1524 (2004).
18. W. Stephan, Y. S. Song, C. H. Langley, The Hitchhiking Effect on Linkage Disequilibrium Between Linked Neutral Loci. *Genetics* **172**, 2647-2663 (2006).
19. H. G. Parker *et al.*, An Expressed Fgf4 Retrogene Is Associated with Breed-Defining Chondrodysplasia in Domestic Dogs. *Science* **325**, 995-998 (2009).
20. E. Cadieu *et al.*, Coat Variation in the Domestic Dog Is Governed by Variants in Three Genes. *Science* **326**, 150-153 (2009).
21. S. R. Browning, B. L. Browning, in *Annual Review of Genetics, Vol 46,* B. L. Bassler, Ed. (Annual Reviews, Palo Alto, 2012), vol. 46, pp. 617-633.

22. I. Zapata, J. A. Serpell, C. E. Alvarez, Genetic mapping of canine fear and aggression. (manuscript submitted for publication).
23. Y. Y. Hsu, J. A. Serpell, Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *Journal of the American Veterinary Medical Association* **223**, 1293-+ (2003).
24. M. Rimbault *et al.*, Derived variants at six genes explain nearly half of size reduction in dog breeds. *Genome Res.* **23**, 1985-1995 (2013).
25. Y. Sun *et al.*, Loss-of-function mutations in IGSF1 cause an X-linked syndrome of central hypothyroidism and testicular enlargement. *Nature Genet.* **44**, 1375-1381 (2012).
26. B. Bai *et al.*, DoGSD: the dog and wolf genome SNP database. *Nucleic Acids Research*, (2014).
27. C. Wu, I. MacLeod, A. I. Su, BioGPS and MyGene.info: organizing online, gene-centric information. *Nucleic Acids Research* **41**, D561-D565 (2013).
28. C. Zancanaro, C. M. Caretta, F. Merigo, A. Cavaggioni, F. Osculati, alpha-gustducin expression in the vomeronasal organ of the mouse. *European Journal of Neuroscience* **11**, 4473-4475 (1999).
29. M. Tizzano, F. Merigo, A. Sbarbati, Evidence of solitary chemosensory cells in a large mammal: the diffuse chemosensory system in Bos taurus airways. *Journal of Anatomy* **209**, 333-337 (2006).
30. O. Dehkordi *et al.*, Neuronal expression of bitter taste receptors and downstream signaling molecules in the rat brainstem. *Brain Research* **1475**, 1-10 (2012).
31. Y.-J. Shin *et al.*, Enhanced Expression of the Sweet Taste Receptors and Alpha-gustducin in Reactive Astrocytes of the Rat Hippocampus Following Ischemic Injury. *Neurochem Res* **35**, 1628-1634 (2010).
32. X. Ren, L. Zhou, R. Terwilliger, S. Newton, I. E. De Araujo, Sweet taste signaling functions as a hypothalamic glucose sensor. *Frontiers in Integrative Neuroscience* **3**, (2009).
33. E. S. Lein *et al.*, Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168-176 (2007).
34. A. Grewal, P. Lambert, J. Stockton, in *Current Protocols in Bioinformatics*. (John Wiley & Sons, Inc., 2002).
35. I. J. Rampil, D. H. Moller, A. H. Bell, Isoflurane modulates genomic expression in rat amygdala. *Anesth. Analg.* **102**, 1431-1438 (2006).
36. K. L. Lunetta *et al.*, Rare coding variants and X-linked loci associated with age at menarche. *Nat Commun* **6**, 7756 (2015).
37. S. D. Joustra *et al.*, Spatial and temporal expression of immunoglobulin superfamily member 1 in the rat. *Journal of Endocrinology* **226**, 181-191 (2015).
38. M. Uhlen *et al.*, Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
39. R. Coppinger, L. Coppinger, *Dogs: a new understanding of canine origin, behavior, and evolution.*, (University of Chicago Press, Chicago, 2001).
40. Y. Li *et al.*, Domestication of the Dog from the Wolf Was Promoted by Enhanced Excitatory Synaptic Plasticity: A Hypothesis. *Genome Biology and Evolution* **6**, 3115-3121 (2014).
41. D. Grimm, Dawn of the dog. *Science* **348**, 274-279 (2015).

42. F. Irie, H. Badie-Mahdavi, Y. Yamaguchi, Autism-like socio-communicative deficits and stereotypies in mice lacking heparan sulfate. *Proceedings of the National Academy of Sciences* **109**, 5052-5056 (2012).
43. A. Yamamoto *et al.*, Structural Abnormalities of Corpus Callosum and Cortical Axonal Tracts Accompanied by Decreased Anxiety-Like Behavior and Lowered Sociability in Spock3-Mutant Mice. *Developmental Neuroscience* **36**, 381-395 (2014).
44. F. C. Zucchi *et al.*, Lifetime stress cumulatively programs brain transcriptome and impedes stroke recovery: benefit of sensory stimulation. *PLoS One* **9**, e92130 (2014).
45. M. Morrow *et al.*, Breed-dependent differences in the onset of fear-related avoidance behavior in puppies. *Journal of Veterinary Behavior: Clinical Applications and Research* **10**, 286–294 (2015).
46. N. H. Dodman *et al.*, Use of fluoxetine to treat dominance aggression in dogs. *J Am Vet Med Assoc* **209**, 1585-1587 (1996).
47. Y. Uchida, N. Dodman, J. DeNapoli, L. Aronson, Characterization and treatment of 20 canine dominance aggression cases. *J Vet Med Sci* **59**, 397-399 (1997).
48. D. J. Anderson, R. Adolphs, A framework for studying emotions across species. *Cell* **157**, 187-200 (2014).
49. E. Karlsson *et al.*, Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B. *Genome Biology* **14**, R132 (2013).
50. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* **44**, 821-824 (2012).
51. S. Purcell *et al.*, PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American journal of human genetics* **81**, 559-575 (2007).
52. M. P. Hoeppner *et al.*, An Improved Canine Genome and a Comprehensive Catalogue of Coding Genes and Non-Coding Transcripts. *PLoS ONE* **9**, e91172 (2014).

Acknowledgments:
We are grateful to the dog owners who contributed behavioral data to C-BARQ, and to Drs. Vaysse, Ratnakumar and Webster and their co-authors and members of the LUPA Consortium, and Drs. Boyko, Bustamante and Ostrander, and their co-authors, for the availability of the rich genotype, mapping and selection data from the Vaysee et al. and Boyko et al. studies cited here. C.E.A. is supported by Department of Defense CDMRP grant award no. W81XWH-11-2-0224, American Kennel Club CHF award no. 01660 and Morris Animal Foundation award no. D13CA-073. The authors declare that they have no conflicts of interest. The data used to conduct the present studies were previously published and are referenced in the text.

Fig. 1: Mapping phased-haplotypes and S_i/D_i regions for known canine GWA loci. (A). Analysis of small-size *IGF1* variation (quantitative trait associated with T allele). (B). Analysis of dog furnishings trait (binary trait associated with T allele). Red line is the peak GWA SNP. Large window sizes are 1.26 Mb in A and 1.48 Mb in B. Small window sizes are 422.3 kb and 439.9 kb, respectively. Green-white gradient shading

represents larger-smaller dog body size in A, and presence/absence of dog furnishings in B. Diagonal lines indicate missing phenotype data.

Fig. 2: Mapping phased-haplotypes and S_i/D_i regions for fear/aggression traits. Quantitative trait with increased-fear/aggression associated with the C allele. Red line is the peak SNP for fear/aggression GWA. The large window size corresponds to 1.50 Mb, and small to 812.1 kb. Green-white gradient shading represents larger-smaller fear/aggression risk. Diagonal lines indicate missing phenotype data.

Fig. 3: Mapping phased-haplotypes and S_i /D_i regions for fear/aggression traits. Quantitative trait with decreased-fear/aggression associated with the G allele. Red line is the peak SNP for fear/aggression GWA, orange is for sociability peak and purple is for size peak. The large window size corresponds to 5.20 Mb, and small to 2.00 Mb. Green-white gradient shading represents larger-smaller fear/aggression risk. Diagonal lines indicate missing phenotype data.

Supplementary Materials:
Fig. S1. Phased haplotype analysis for two loci of unknown selection status
Fig. S2. Mouse Gnat3 is highly expressed in the Amygdala and Piriform area

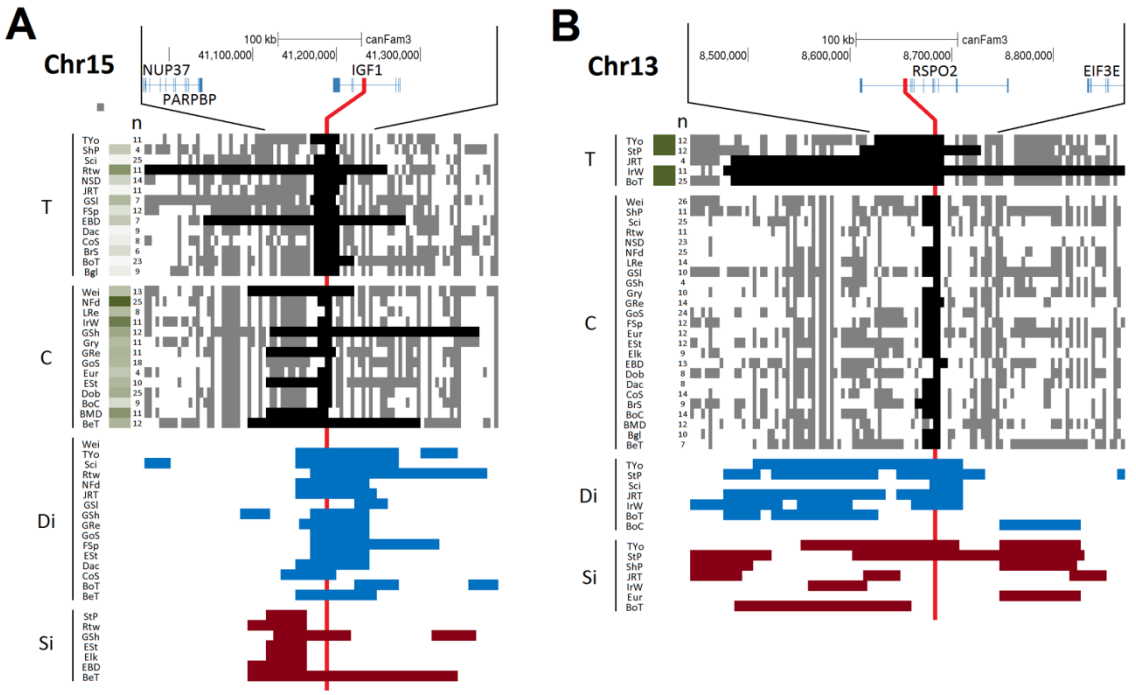

A
Chr15
100 kb
canFam3
41,100,000
41,200,000
41,300,000
NUP37
PARPBP
IGF1
n
T
C
Di
Si
B
Chr13
100 kb
canFam3
8,500,000
8,600,000
8,700,000
8,800,000
RSPO2
EIF3E
n
T
C
Di
Si

Fig. 1

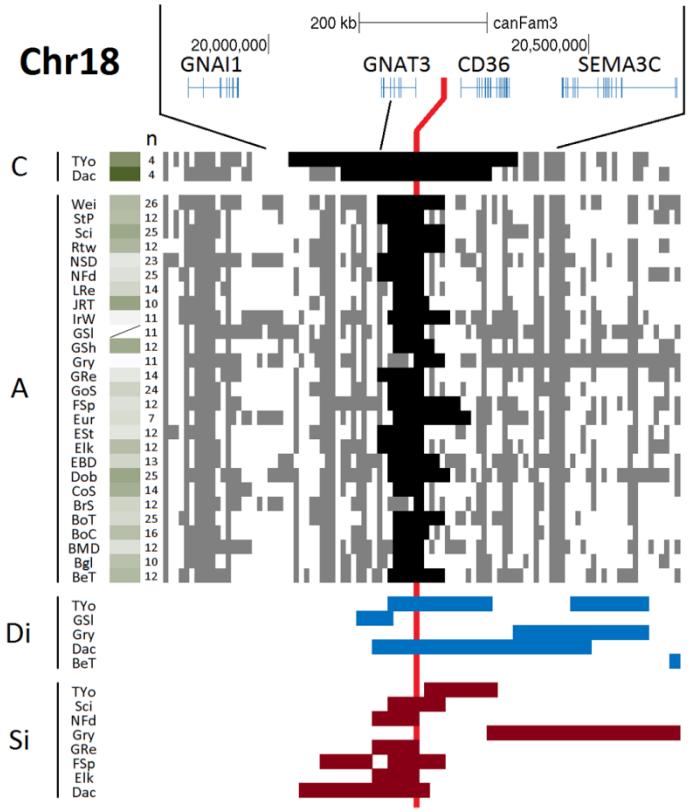

Chr18
200 kb
canFam3
20,000,000
20,500,000
GNAI1
GNAT3
CD36
SEMA3C
n
C
TYo 4
Dac 4
A
Wei 26
StP 12
Sci 25
Rtw 12
NSD 23
NFd 25
LRe 14
JRT 10
IrW 11
GSl 11
GSh 12
Gry 11
GRe 14
GoS 24
FSp 12
Eur 7
ESt 12
Elk 12
EBD 13
Dob 25
CoS 14
BrS 12
BoT 25
BoC 16
BMD 12
Bgl 10
BeT 12
Di
TYo
GSl
Gry
Dac
BeT
Si
TYo
Sci
NFd
Gry
GRe
FSp
Elk
Dac

Fig. 2

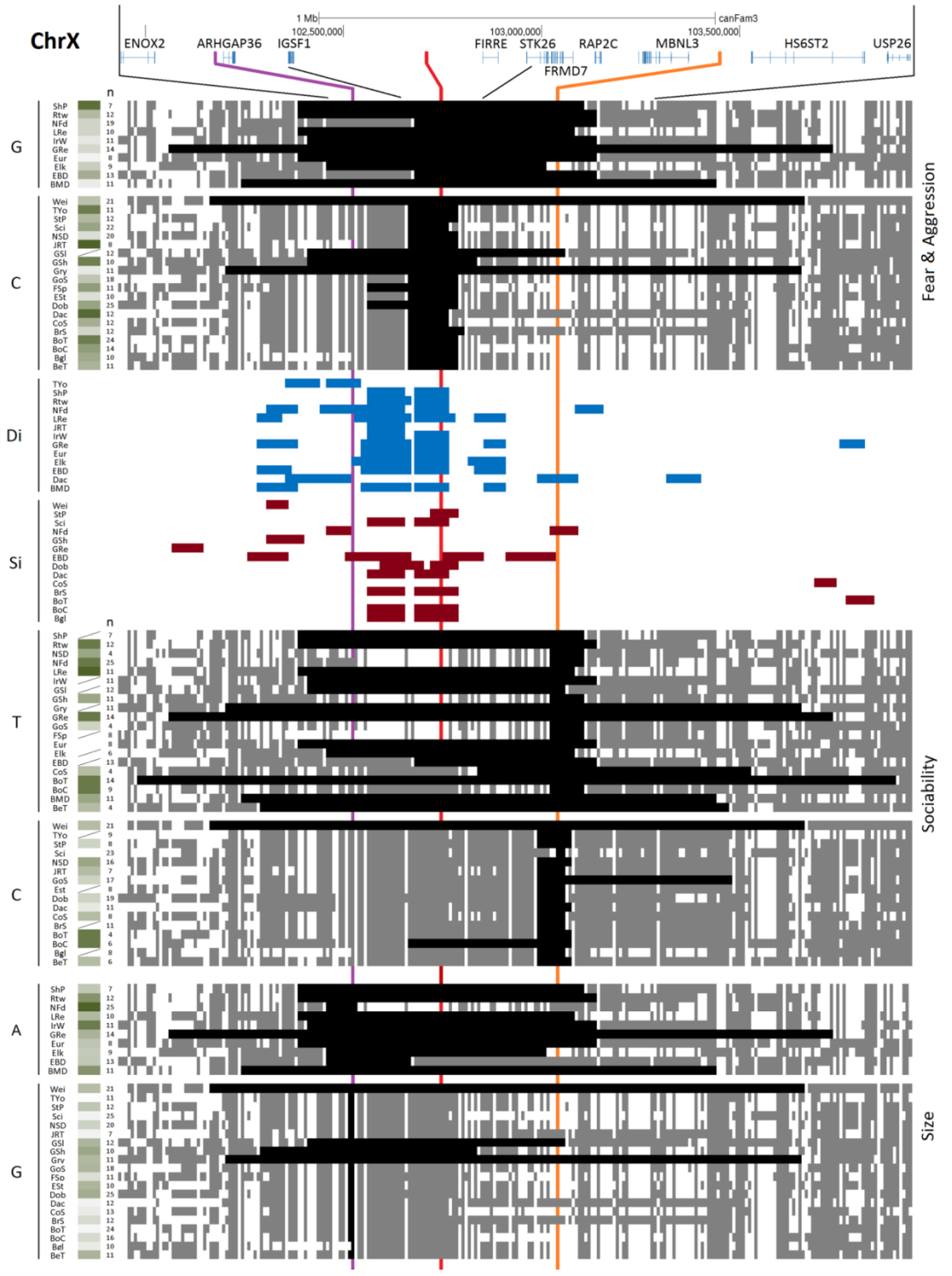

ChrX
1 Mb
canFam3
102,500,000
103,000,000
103,500,000
ENOX2
ARHGAP36
IGSF1
FIRRE
STK26
RAP2C
MBNL3
HS6ST2
USP26
FRMD7
n
G
C
Di
Si
T
C
A
G
Fear & Aggression
Sociability
Size

Fig. 3.

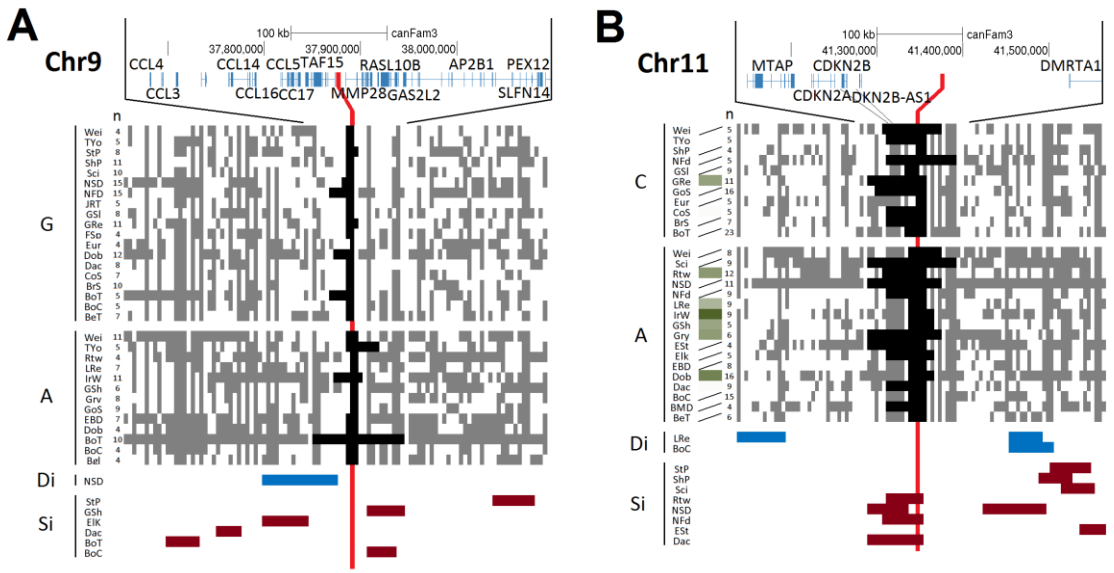


Fig. S1. Phased haplotype analysis for two loci of unknown selection status. (A) Region not associated with any known trait. (B) Dog osteosarcoma risk locus (binary trait). Large window sizes: 1.60 Mbp and 1.48 Mbp respectively. Small window sizes: 441.2 kb and 424.3 kb respectively. Green-white gradient shading represents percent of incidence of dog osteosarcoma according to literature. In (A), although the general patterns are unremarkable, the salient feature is that the largest phased haplotype (in black) is in the Border Terrier, and that is flanked by reduced variation (gray). This suggested the possibility that the highlighted haplotype could be in LD with a larger phased haplotype. In this case, the alternative allele is also present in the same breed and comparison of the two patterns shows a shared segment of reduced variation, which overlaps an S_i interval unique to the Border Terrier. We tested this possibility and confirmed that there is a 391 kb phased haplotype that overlaps the S_i region, and that is fixed in the Border Terrier; further investigation is necessary to consider the evolutionary basis for this observation. We also briefly considered the possibility that complex diseases that are common in dogs could be the result of selection for unrelated traits. We chose the locus associated with osteosarcoma risk (B) in Greyhounds because the risk allele is common across breeds and was shown to be fixed in two other breeds with high risk of this bone cancer, the Rottweiler and Irish Wolfhound [1]. That peak SNP in the *CDNK2A/B* locus on chr11 (1.45 Odds Ratio) intersects S_i regions in three breeds, two of those have high risk of osteosarcoma and the third does not (another has an S_i region shifted ~10kb upstream); all four of those S_i regions carry the osteosarcoma risk allele with the exception of the Newfoundland when both alleles are present. Phasing analysis is inconclusive because both the risk and alternative alleles in several breeds have haplotypes in the range of 50-100kb. However, it may be of interest that the minimal region of overlap across all risk haplotypes is significantly larger than the alternative haplotype despite the former being more common across breeds (37 vs. 21kb). This hints at the possibility that the risk allele was under selection long ago. The *CDKN2A/B/AS1* locus is genetically and biologically very complex, with central roles in cell cycle control that affect the balance between stemness and senescence. Those coding and non-coding genes are among the most highly mutated in human cancer, but, in their

wild type forms, are also critical to diverse other aspects of human health. Based on the broad and key biological relevance, it seems likely that there was natural selection for several variants at this locus prior to breed creation. Some of those could be under purposeful or incidental human-selection in modern breeds (e.g., histiocytic sarcoma in Bernese Mountain Dogs [2]). It is thus notable that two breeds exhibit reduced variation in the haplotype phasing analysis (i.e., the top length of haplotype homozygosity shown in black, flanked by regions of reduced variation in gray), and both have S_i signal in the region.

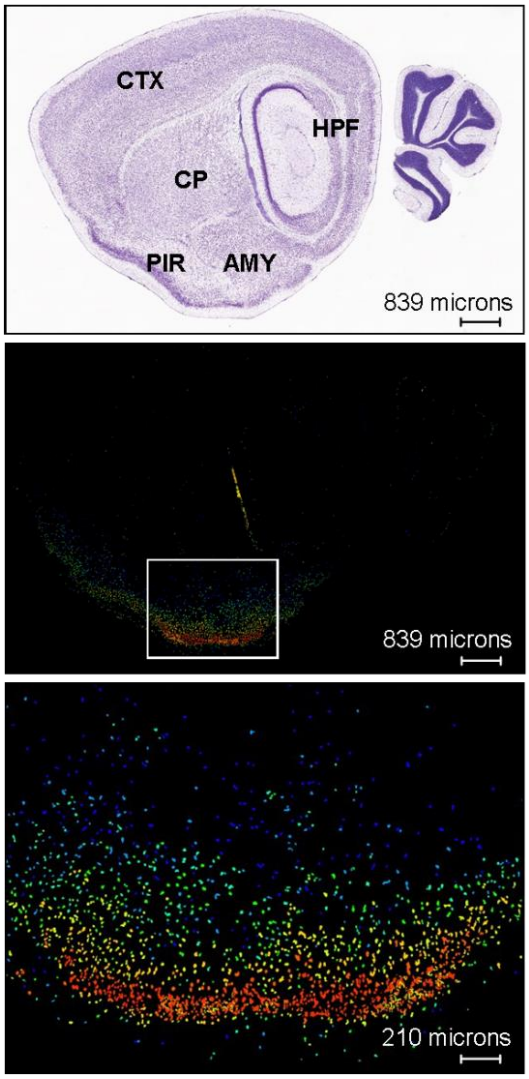


Fig. S2. Mouse *Gnat3* is highly expressed in the Amygdala and Piriform area. Analysis of the Allen Brain Atlas [3] of *in situ* mRNA hybridization data revealed the highest levels of *Gnat3* mRNA expression in the adult mouse to be in the Amygdala and Piriform area (AMY and PYR). The top image is Nissl staining and the bottom two are *in situ* hybridization; signal intensity is represented in a blue (low) to red (high) color spectrum (Atlas Gnat3 Image 1 of 19). The strongest signal is in layer 2 of the Cortical amygdalar area and of the Piriform area. The closeup image at the bottom shows two regions where the most intense color corresponds to the PIR layer 2 (left area of solid red cells) and COA layer 2 (right area of solid red cells); these are separated by an area of intermediate signal strength (orange and yellow cells). The other signal above and perpendicular to AMY/PIR corresponds to the choroid plexus of the lateral ventricle. Another brain area with robust signal is the pons, specifically, in the Lateral reticular nucleus (mixture of

red and orange-yellow cells), Paragigantocellular reticular nucleus, lateral part (yellow and blue cells), and the Facial motor nucleus (yellow and blue cells) (see Atlas Images 10 and 13; data not shown). There is also weak to moderate signal in the mitral layer of the olfactory bulb. The Atlas series has uneven results, with most sections having very low signal. It is thus possible that some brain areas with detectable expression were missed in these experiments. However, the positive pattern shown in the closeup here is not suggestive of non-specific binding. For support of the image shown here, see also Atlas Image 5.

References

1. Karlsson, E., et al., *Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B.* Genome Biology, 2013. **14**(12): p. R132.
2. Shearin, A.L., et al., *The MTAP-CDKN2A Locus Confers Susceptibility to a Naturally Occurring Canine Cancer.* Cancer Epidemiology Biomarkers & Prevention, 2012. **21**(7): p. 1019-1027.
3. Lein, E.S., et al., *Genome-wide atlas of gene expression in the adult mouse brain.* Nature, 2007. **445**(7124): p. 168-176.

Appendix 7

Published article: Human Genetic Relevance and Potent Antitumor Activity of Heat Shock Protein 90 Inhibition in Canine Lung Adenocarcinoma Cell Lines

Clemente-Vicario F, Alvarez CE, Rowell JL, Roy S, London CA, Kisseberth WC, Lorch G. Human Genetic Relevance and Potent Antitumor Activity of Heat Shock Protein 90 Inhibition in Canine Lung Adenocarcinoma Cell Lines. PLoS One. 2015 Nov 11;10(11):e0142007. doi: 10.1371/journal.pone.0142007. eCollection 2015. PubMed PMID: 26560147; PubMed Central PMCID: PMC4641597.

PLOS | ONE

RESEARCH ARTICLE

Human Genetic Relevance and Potent Antitumor Activity of Heat Shock Protein 90 Inhibition in Canine Lung Adenocarcinoma Cell Lines

Francisco Clemente-Vicario[1], Carlos E. Alvarez[2], Jennie L. Rowell[3], Satavisha Roy[1], Cheryl A. London[4], William C. Kisseberth[1], Gwendolen Lorch[1]*

1 Department of Veterinary Clinical Sciences, College of Veterinary Medicine, The Ohio State University, Columbus, Ohio, United States of America, **2** Center for Molecular and Human Genetics, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio, United States of America, **3** College of Nursing, The Ohio State University, Columbus, Ohio, United States of America, **4** Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, Columbus, Ohio, United States of America

* lorch.2@osu.edu

OPEN ACCESS

Citation: Clemente-Vicario F, Alvarez CE, Rowell JL, Roy S, London CA, Kisseberth WC, et al. (2015) Human Genetic Relevance and Potent Antitumor Activity of Heat Shock Protein 90 Inhibition in Canine Lung Adenocarcinoma Cell Lines. PLoS ONE 10(11): e0142007. doi:10.1371/journal.pone.0142007

Editor: Gabriele Multhoff, Technische Universitaet Muenchen, GERMANY

Received: July 22, 2015

Accepted: October 15, 2015

Published: November 11, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: These studies were funded by the National Canine Cancer Foundation; Grant GL15OH-005 to GL (http://www.wearethecure.org/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interest exist.

Abbreviations: HSP90, heat shock protein ninety; HSP70, heat shock protein seventy; EGFR,

Abstract

Background

It has been an open question how similar human and canine lung cancers are. This has major implications in availability of human treatments for dogs and in establishing translational models to test new therapies in pet dogs. The prognosis for canine advanced lung cancer is poor and new treatments are needed. Heat shock protein 90 (HSP90) is an ATPase-dependent molecular chaperone ubiquitously expressed in eukaryotic cells. HSP90 is essential for posttranslational conformational maturation and stability of client proteins including protein kinases and transcription factors, many of which are important for the proliferation and survival of cancer cells. We investigated the activity of STA-1474, a HSP90 inhibitor, in two canine lung cancer cell lines, BACA and CLAC.

Results

Comparative genomic hybridization analysis of both cell lines revealed genetic relevance to human non-small cell lung cancer. STA-1474 inhibited growth and induced apoptosis of both cell lines in a dose- and time-dependent manner. The IC_{50}s after 72 h treatment with STA-1474 were 0.08 and 0.11 μM for BACA and CLAC, respectively. When grown as spheroids, the IC_{50} of STA-1474 for BACA cells was approximately two-fold higher than when grown as a monolayer (0.348 μM *vs*. 0.168 μM), whereas CLAC spheroids were relatively drug resistant. Treatment of tumor-stromal fibroblasts with STA-1474 resulted in a dose-dependent decrease in their relative cell viability with a low IC_{50} of 0.28 μM.

epidermal growth factor receptor; c-Kit, v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog; HER2, human epidermal growth factor receptor 2; VEGFR, vascular endothelial growth factor receptor; PDGFRα/β, platelet-derived growth factor receptor alpha and beta; c-MET, c-mesenchymal epithelial transition factor; MAPK, mitogen-activated protein kinases; c-RET, c-rearranged during transfection; AKT-2, v-akt murine thymoma viral oncogene homolog 2; NKX2-1, NK2 homebox 1; TTF-1, thyroid transcription factor 1; IGF-IRβ, insulin growth factor I; STAT3, signal transducer and activator of transcription 3; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; NDUFA1, NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1; ALK, anaplastic lymphoma kinase; FACS, fluorescence-activated cell sorting; CI, combination index; DRI, dose reduction index; FGFR3, fibroblast growth factor receptor 3; CAFs, cancer associated fibroblasts; CGH, comparative genomic hybridization; CNAs, copy number alterations; CaGe, cancer gene annotation system for cancer genomics.

Conclusions

Here we first established that lung adenocarcinoma in people and dogs are genetically and biochemically similar. STA1474 demonstrated biological activity in both canine lung cancer cell lines and tumor-stromal fibroblasts. As significant decreases in relative cell viability can be achieved with nanomolar concentrations of STA-1474, investigation into the clinical efficacy of this drug in canine lung cancer patients is warranted.

Introduction

Cancer is the leading cause of mortality in dogs [1–3]. Lung cancer has an incidence of 15/100,000 dogs per year [4] and it is generally a disease of older dogs with an approximate age of 11 years old at the time of diagnosis. The most common histological subtype of canine lung cancer is adenocarcinoma, representing 74–77% of cases [5, 6]. Less than one-third of cases have localized disease, with 23% having distant metastasis, 13.5% lymph node metastasis and 34.6% diagnosed with vascular/lymphatic or intrapulmonary spread [7]. Clinical staging is an essential requirement for determining prognosis and treatment and surgical excision is the most commonly used and effective treatment modality. Dogs with solitary tumors (T1) have a median survival time after surgery of 348 days, while those that have multiple tumors of any size (T2) or tumors invading neighboring tissues (T3), have a dismal median survival time of only 58 days after a lung lobectomy [8].

The lack of good treatment options, other than surgery, has driven us to improve our understanding of the molecular basis of this cancer in dogs, in order to enhance the development of novel and rational therapies for this disease. Currently selected nonsurgical treatment options for dogs with lung cancer are based on clinical acumen, drug—treatment responses noted in other adenocarcinoma tumor subtypes, extrapolation of drug responses in people, and a few drug toxicity studies that evaluated a very limited number of dogs with advanced lung cancer. Therefore, a significant need exists to provide a rationale for treatment selection based on more robust evidence from *in vitro* and *in vivo* models of canine lung cancer [9].

Lung cancer remains the most common cause of cancer-related mortality in people. People with advanced disease are treated with medical therapy alone and have a poor prognosis with an overall five-year survival less than 15% [10]. Discovery of a spectrum of gene mutations and genomic aberrations has led to the use of targeted therapies utilizing a precision medicine approach which has been associated with often dramatic, although often short-lived, clinical benefit [11, 12]. Unfortunately, even in patients treated with first-line targeted therapy, resistance invariably develops, leaving chemotherapy as the cornerstone of subsequent therapy [13]. Pet dog translational models represent a major opportunity to better understand and treat human cancers, but lung cancer is the most common human cancer yet to be genetically dissected in dogs [14]. Because dog breeds are on the order of 100-fold more genetically simple than the human or dog populations, they are more powerful for understanding germline-genetic, environmental and gene-gene interaction risks [14]. Notably, the availability of state of the art human treatments for canine lung cancer is also dependent on this knowledge.

Heat shock protein 90 (HSP90), a molecular chaperone protein, plays a central role in regulating the folding, stability and function of many proteins that are oncogenic drivers for lung cancers. HSP90 is a highly conserved protein that folds newly synthesized proteins into their biologically active conformations preventing their aggregation. HSP90 is expressed as a 90 kDa protein with two major isoforms (HSP90α and HSP90β) and plays an essential role in

maintaining cellular protein homeostasis. Co-chaperones and client proteins can modify HSP90's mechanism of action [15–17]. Tumor cells express high levels of HSP90, which exists in highly activated complexes that are particularly susceptible to binding HSP90 inhibitors [18]. Heat-shock proteins promote tumor cell survival, growth and metastasis, even in growth-factor deprived conditions, by allowing continued protein translation and cellular proliferation [19]. These proteins provide a mechanism whereby cellular stresses experienced by cancer cells are either managed or avoided. Many oncogenes, including tyrosine kinases, transcription factors and cell-cycle regulatory proteins are clients of HSP90, and thus HSP90 is recognized as a crucial facilitator of cancer cell survival [20, 21].

Pharmacological blockade of HSP90, i.e. HSP90 inhibition, represents an alternative approach for therapeutic intervention, and has shown efficacy in both preclinical studies and clinical trials in people [22–24]. Geldamycin, a benzoquinone ansamycin antibiotic, binds to the nucleotide-binding site of the N-terminal domain of HSP90 preventing ATP binding, resulting in HSP90 inhibition. Geldamycin has poor solubility, stability and unacceptable liver toxicity in dogs at therapeutic doses therefore, analogues were developed [25]. STA-1474 is a highly soluble prodrug of ganetespib, a novel resorcinol-containing compound unrelated to geldamycin that binds in the ATP-binding domain at the N-terminus of HSP90 and acts as a potent HSP90 inhibitor. A phase I study with STA-1474 in dogs with cancer showed clinical activity with low grade gastrointestinal toxicity that was manageable with concomitant medications [26]. Inhibiting HSP90 in lung cancer is appealing as no resistance mutations have been identified, suggesting it represents a relatively stable target for drug treatment. As little is known about the efficacy of cytotoxic and small molecule inhibitors in canine lung cancer, the purpose of this study was to characterize the activity of currently used chemotherapeutic agents and the small molecule inhibitors, torceranib phosphate, crizotinib and STA-1474 and the effects of HSP90 inhibition on the mRNA expression of relevant kinases and HSP90 client proteins in two canine lung cancer cell lines. Here we show that STA1474 demonstrated biological activity in both canine lung cancer cell lines and tumor-stromal fibroblasts.

Materials and Methods

Cell Lines and Reagents

The BACA cell line was generously provided by Dr. Joseph J. Wakshlag, Cornell University College of Veterinary Medicine (Ithaca, NY). The BACA cell line was established from a histologically confirmed canine primary lung adenocarcinoma. Immunostaining of the cell line was positive for cytokeratin indicating epithelial origin [27]. The CLAC cell line was purchased through an approved materials transfer agreement with the Japan Health Sciences Foundations, JCRB Cell Bank (Osaka, Japan) [28]. Both cell lines were maintained in high-glucose Dulbecco modified Eagle medium (DMEM) with GlutaMax (Invitrogen, Carlsbad, CA) and supplemented with 10% heat-inactivated fetal bovine serum (FBS), and a penicillin- (100 I.U./ml) streptomycin (100 μg/ml) solution. Cells were passaged at ~90% confluence. *In vitro* experiments were performed when cells were ~90% confluent. STA-1474 was kindly provided by Synta Pharmaceuticals® (Lexington, MA). Crizotinib and toceranib were provided by Zoetis™ (Groton, CT). Carboplatin (Teva Pharmaceuticals Ltd, Sellersville, PA), gemcitabine (Accord Healthcare, Durham, NC) and vinorelbine (Mylan Institutional, Rockford, IL) were purchased from The Ohio State University Veterinary Medical Center Pharmacy (Columbus, OH).

Comparative Genomic Hybridization Array

We custom designed a 966,903 feature comparative genomic hybridization (CGH) array tiling the canine genome (C.E.A and J.L.R, manuscript in preparation; see also [29, 30]. The array is

comprised of isothermal 60-nucleotide probes targeting all regions of unique or low copy repeats (i.e., with otherwise-unique sequence) based on the CanFam2 assembly (including the unmapped contigs annotated as chrUn). Average spacing of probes is 1.9kb for unique sequence and 1.2kb for low copy repeats. The CGH platform and probe design method is Agilent SurePrint G3 (it includes 7,113 additional Agilent control probes). DNA quality control, array hybridization and scanning were performed by Asuragen® (Austin, TX) under Agilent certified conditions. The two tumor cell line samples were compared against a healthy male Labrador retriever as the reference. The reference specimen was obtained under informed owner consent and the following Ohio State University IACUC approved protocol (2010A0015-R1, Canine Specimen Collection and Banking) which covered the procedure used to obtain the sample and their subsequent use for research application. Agilent uses a linear normalization process (including dye based normalization using copy-neutral normalization probes) for their LogR values. This data was imported into Golden Helix SNP and Variation Suite, and converted from LogR10 to LogR2. Sample quality metrics were performed, including percentile based Winsorizing, derivative log ratio spread, and wave detection/correction. For segmentation, we used the univariate Optimal Copy Number Analysis Module (CNAM in Golden Helix), which uses a change-point identification algorithm. While this powerful algorithm accurately identifies changes in sequential data, it is computationally intensive. We selected 10 max segments per 10,000 bases, 20-marker minimum for a copy number call, and a max pairwise permuted *p*-value of 0.005 (with 2000 permutations per pair). CNV calls were based on a logR2 ratio threshold of -0.40/0.40 for losses and gains, respectively (Excel worksheet #1, S1 Table). Because this segmentation algorithm is more sensitive to deletion copy number changes, we used a more stringent threshold for deletions for further analysis (-0.45 for deletions and 0.4 for gains). Mean symmetric smoothing was applied to all figures.

RNA Isolation and Reverse-Transcriptase Polymerase Chain Reaction

Total RNA was isolated from both cell lines using the Absolutely RNA Miniprep Kit (Agilent Technologies, La Jolla, CA), according to manufacturer instructions. The RNA quantity and quality was assessed with a NanoVue Spectrophotometer (GE Healthcare, Piscataway, NJ). TaqMan® Reverse Transcription Kits (Applied Biosystems, Carlsbad, CA) were used to make cDNA for RT-PCR analysis of all gene transcripts. Primer sequences used for the RT-PCR reactions are listed in Table 1. Additional primer pair sequences used for RT-PCR reactions are referenced in Mariotti *et al.* [31]. The amplified cDNA products were separated according to size using gel electrophoresis. Amplicons were resolved on a 1% agarose gel to visualize the products. NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1 (NDUFA1) was used as a housekeeping gene [32].

Sequencing and Sequence Alignment

Standard PCR was used to generate high fidelity *Taq* polymerase-amplified PCR products. The resolved PCR products were extracted from the gel, purified using QIAquick PCR Purification Kit (Qiagen, Germantown, MD) and sequenced using BigDye™ Terminator Cycle Sequencing chemistry (Applied Biosystems, Carlsbad, CA). Automated Sanger capillary sequencing reactions for were run on a 3730 DNA Analyzer. Sequence alignments for the genes *HSP70*, *HSP90AA1*, *HSP90AB1*, *HSP90B1*, *MET*, *NDUFA1* and *NKX2-1* were made to the reference sequences NCBI: [*Canis lupus familiaris* (dog)]Gene ID: 403612, 480438, 474919, 404019, 403438, 481033, and 403940 updated on 7-Dec-2014 and 29-Jan-2015) using the ClustalW procedure in DNASTAR Software Lasergene MegAlign® v.12.1 Madison, WI. All other primer

Table 1. Canine RT-PCR Primers.

Gene	Forward	Reverse	Product size (bp)
AKT	5'-TGCTTAAGAAGGACCCCAAGC-3'	5'-GCTGGTCCAGTTCGAGGGA-3'	**253**
ALK	5'-CTGTATCGGGGTGAGTCTGC-3'	5'-CAGGGCCTGGACAGGTTAAG-3'	**245**
c-KIT	5'-GATGGCCCCTGAGAGCATTT-3'	5'-GCCTTTTCAGGGGATCAGCA-3'	**248**
c-MET	5'-GAGGAATGTCCCACTGGAGC-3'	5'-TGCTGTCCCTCGACCATTTG-3'	**297**
EGFR	5'-TGGTCCTGGGGAATTTGGAA-3'	5'-GGTTATTGCTGAAGCGCACA-3'	**289**
ErbB2/HER2	5'-CCCGAGACCCACCTGGATA-3'	5'-CAGGGCGTAGTTGTCCTCAA-3'	**228**
HSP70	5'-AGCTGGAGCAGGTGTGTAAC-3'	5'-GGGGAAGAAGTCCTAATCCACC-3'	**146**
HSP90AA1	5'-ACCGAACTGGCTGAAGACAA-3'	5'-GATCACTTCCAGGCCATGCT-3'	**285**
HSP90AB1	5'-TGTCCGCCGTGTGTTTATCA-3'	5'-GCGCCGGTTAGTGGAATCTT-3'	**280**
HSP90B1	5'-TGAAACTGTTGAGGAGCCCA-3'	5'-GTGACTTCCCCTTCAGCAGT-3'	**288**
NDUFA1	5'-AATATTATAAATGGGAGGCGCGG-3'	5'-TAGTGAACCTGTGGATGTGCG-3'	**168**
NKX2-1/TTF-1	5'-GACGTGAGCAAGAACATGGC-3'	5'-CAGATTTTGACCTGCGTGGG-3'	**298**
PDGFRα	5'-CGACTCCGTTCTCAGTGTCT-3'	5'-CTGTTCCGCATGGTGTCCT-3'	**223**
PDGFRβ	5'-CCACGCCTCTGACGAGATTT-3'	5'-CTGCACAGCAGTGTACAGGA-3	**259**
RET	5'-AGCAGGATACACGACCGTTG-3'	5'-GTCCCGATGGACAAGCTTCA-3'	**384**

doi:10.1371/journal.pone.0142007.t001

sets used to generate PCR products were previously sequenced and aligned to verify the amplicons [31]

Cell Proliferation Assay

To assess relative cell proliferation, cells were seeded in 96-well plates in 100 μl of DMEM supplemented with 10% FBS and incubated overnight. 2.5 x 10^3 and 4 x 10^3 cells per well were seeded for CLAC and BACA, respectively in order to achieve ~ 90% confluency. Plates were then treated with increasing concentrations of gemcitabine, vinorelbine, carboplatin, crizotinib, toceranib phosphate or STA-1474 and were evaluated after 72 h, using the CyQUANT® cell proliferation assay according to the manufacturer instructions (Molecular Probes, Eugene, OR). For each drug and concentration, six wells were used. Briefly, 72 h after treatment, media was removed by gently inverting the plates and the plates were frozen at -80°C. The following day, plates were thawed at room temperature (RT) and 200 μl of CyQUANT® GR dye/cell-lysis buffer was added to each well. Plates were incubated at RT for 5 min, protected from light and then fluorescence measurements were made using a plate reader (Molecular Devices, Sunnyvale, CA), with excitation at 485 nm and emission detection at 530 nm. Relative cell number was calculated as a percentage of the control wells: absorbance of sample/absorbance of DMSO treated cells x 100.All proliferation experiments were repeated three times.

Similarly, 4 x 10^3 cells of BACA and CLAC were seeded per well and incubated overnight before a 72 h treatment with increasing concentrations of VER155008 and STA-1474. Cell proliferation was evaluated with CyQUANT® cell proliferation assay as described above. The drug concentrations were selected based on predetermined half maximal concentration 50% (IC_{50}) values for each drug. For the VER155008 and STA-1474 studies, fixed constant ratio drug combinations from 0.0625 to 16*X*, where *X* is the IC_{50} were evaluated using a minimum of 9 data points which were each repeated in triplicate. The combinations were evaluated for synergism, additive effects or antagonism by median-effect analysis (CompuSyn Software v. 1, Inc, Paramus, NJ) [33]. The nature of the combinatorial interactions was evaluated using the combination index (CI) method. Briefly, cytotoxic effects of the drug combination are described by the equation $f_a/f_u = [D/D_m]^m$ where f_a is the fraction of cells affected, f_u is the fraction of cells

not affected (1-f_a), D is the dose of drug, D_m is the dose of drug to cause the median effect and m is the slope of the median-effect curve. The CI value definitions are listed in the Table 2.

Detection of Apoptosis

Alexa Fluor 488 annexin V and 1 μl of PI solution (100 μg/ml) were used to stain cells for fluorescence-activated cell sorting (FACS). After 15 min of incubation at RT, 400 μl of annexin-binding buffer was added and the sample was gently mixed and kept on ice until analysis. Cells were analyzed within 30 min of staining. Caspase 3/7 activity was evaluated with the Senso-Lyte® Homogeneous AMC Caspase 3/7 Assay Kit (AnaSpec, Fremont, CA) according to the manufacturer's instructions. Briefly, 4,000 cells per well were seeded in 100 μl of medium and incubated overnight. The next morning, cells were treated with STA-1474 (0.05–1 μM) for 24 and 48 h. Then, 50 μl of the caspase 3/7 substrate solution was added to each well and mixed in a plate shaker for 60 min at 150 rpm and wrapped in foil to protect from direct light. Fluorescence intensity was measured using a plate reader (Molecular Devices), with excitation at 245 nm and emission detection at 442 nm.

Protein Isolation

For the measurement of HSP90 and HSP70, client proteins and phosphorylated forms in the BACA and CLAC treated cells, cells were grown to 90% confluence in 100-mm dishes, placed on ice, rinsed with ice-cold DPBS, and lysed with 1X cell lysis buffer (#9803, Cell Signaling Technology®, Danvers, MA) with 1mM phenylmethanesulfonyl fluoride and protease inhibitors (Halt Protease Inhibitor Cocktail Kit; Pierce, Rockford, IL) added just before use. Cells were scraped from the dishes, and the lysates were incubated in the buffer for 15 min on ice, and centrifuged for 20 min at 16,000 x *g* at 4°C. The supernatants were collected, and protein concentration was determined by a modified Bradford method (Bio-Rad Laboratories, Inc, Hercules, CA).

Immunoblot Analyses

For immunoblot studies, BACA and CLAC cells were seeded in triplicate at a density of 3.5 x 10^6 cells per 100 mm plate and incubated overnight. All immunoblot analyses represent protein expression after 72 h of treatment. Cells were treated with 0.05, 0.25, 0.75, and 1 μM of toceranib phosphate and 0.005, 0.05, 0.5 and 1 μM of STA-1474 and were incubated for 72 h. For immunoblot analysis, 3X Laemmli sample buffer with 1 mM β-mercaptoethanol was added to 40 μg of protein

Table 2. Combination index (CI) value ranges with verbal descriptors.

CI	*Description*
<0.10	Very Strong Synergism
0.10–0.30	Strong Synergism
0.30–0.70	Synergism
0.70–0.85	Moderate Synergism
0.85–0.90	Slight Synergism
0.90–1.10	Nearly Additive
1.10–1.20	Slight Antagonism
1.20–1.45	Moderate Antagonism
1.45–3.30	Antagonism
3.30–10.0	Strong Antagonism
>10.0	Very Strong Antagonism

doi:10.1371/journal.pone.0142007.t002

extract (35–200 μg) at a final concentration of 33%, and the samples were heated at 100°C for 5 min. Cell protein extracts were fractioned on a precast 12% NuPAGE® Novex® Bis-Tris polyacrylamide Mini gels (10 cm x 10 cm) using NuPAGE® MES SDS running buffer for small proteins (2–200 kDa) or NuPAGE® MOPS SDS running buffer (Life Technologies, Grand Island, NY) for medium-size proteins (14–200 kDa) followed by electrophoretic transfer to nitrocellulose membranes (Pall Life Sciences, Ann Arbor, MI). The membranes were incubated overnight at 4°C with 0.2% Tween-20 in TBS and 2% bovine serum albumin with rabbit monoclonal and polyclonal antibodies. The blots were incubated with secondary anti-rabbit IgG horseradish peroxidase—linked antibody in 0.2% Tween-20 in TBS and 2% nonfat dry milk for 1 h at RT. Primary antibodies were epidermal growth factor receptor (EGFR, Santa Cruz Biotechnology, Inc., #sc-03; 1:500)[34], STAT3 (antibody #610189, BD Transduction Biosciences, Franklin Lakes, NJ, 1:2000), HSP90 (ADI-SPA-835, 1:600), HSP70/HSP72 (ADI-SPA-810, Enzo Life Sciences, Inc, Farmingdale, NY, 1:2000) insulin growth factor I (IGF-IRβ) receptor beta (#3027, 1:500), HER2 (#4290S, 1:1000), phospho-AKT (#4060P, 1:2000), mTOR (#2972S, 1:1000), phospho-mTOR (#2971S, 1:1000), MAPK (#4695P, 1:1000), phospho-MAPK (4060P, 1:1000), phospho-STAT3 (#9134P, 1:250) phospho-S6 ribosomal protein (#2211S 1:1000), S6 ribosomal protein (#2217S, 1:1000), β-actin (#4970S, 1:1000), GAPDH (#5174P, Cell Signaling Technologies®, Danvers, MA, 1:1000)[35]. Two wells of each pre-cast gel were loaded with markers, one for protein electrophoresis transfer from gel to membrane marker (#10748–010, BenchMark™ Pre-stained protein ladder, Life Technologies) and a protein standard marker (#LC5602, MagicMark™ XP Standard, Life Technologies).

Spheroid and Fibroblast Proliferation Assays

Self-assembled clusters of cell colonies cultured in a microenvironment where cell-cell interactions dominate over cell-substrate interactions were generated in the form of 3-D spheroids. Monolayer cultures were plated at the same time for comparison of treatment effects. In both cases, 5,000 cells per well were seeded and allowed to grow for three days. For the spheroid growth, ultra-low attachment plates (Corning® Costar® Ultra-low attachment multiwall plates, Sigma-Aldrich, St. Louis, MO) were used. For the monolayer model, cells were plated in 96-well black flat-clear bottom plates (Greiner Bio-One GmbH, Frickenhausen, Germany). The wells that formed the perimeter of the 96-well plate (two outer rows) were filled with PBS to minimize edge effect. After 72 h of growth, cells were treated with DMSO or 0.005–10 μM of STA-1474 for 72 h. Then, viability was assessed with the CellTiter-Glo® Luminescent Cell Viability Assay (Promega, Madison, WI) according to manufacturer instructions. Briefly, spheroids were disrupted and mixed by repeated pipetting, aspirated and transferred into a well of a black, flat-clear bottomed multiwell plate in 100 μl of media. Next, 100 μl of CellTiter-Glo® reagent was added to each well. Contents were mixed for 2 min using a shaker to induce lysis and plates were incubated at RT for 10 min before luminescence was recorded using a plate reader (Molecular Devices). Finally, luminescence was normalized to the control group and $IC_{s_{50}}$ were calculated for each cell line and growth model.

For the fibroblast assay, 1.2 x 10^4 cells were seeded per well in 500 μl of medium, incubated for 24 h and then treated with STA-1474 (0.001–1 μM) for 72 h. Viability was assessed with the CyQUANT® Assay (Molecular Probes) according to the manufacturer instructions. Fluorescence was measured using a plate reader (Molecular Devices), with excitation at 485 nm and emission detection at 530 nm and results were normalized to the control group.

Statistical Analyses

Experiments were performed three times and the data presented as mean values ± SD. Statistical analysis of significance was performed using the One-Way ANOVA followed by Bonferroni

test. For non-normal distribution, a Kruskal-Wallis test followed by Dunns test was used. IC_{50} calculations were made using a logarithmic regression curve with Prism® 5 for Mac OS X (GraphPad Software, Inc. La Jolla, CA). *P*-values <0.05 were considered statistically significant.

Results

Structural variation in the BACA and CLAC cell lines authenticates their relevance as a comparative oncogenomic model for human NSCLC

Comparative genomic hybridization was conducted on BACA [27] and CLAC [28] to establish the genes affected by genomic alterations or Copy Number Alterations (CNAs; Fig 1A) [14, 29, 30]. Stringent thresholds (0.4 and 0.45 Log2 ratios) and high minimum-number of probes per CNA segment (20 probes) were applied. All CNAs are provided in S1 Table, segregated according to cell line and size (focal, defined as <3Mb, and large). Many large alterations affect known cancer driver genes in these cell lines. For example, chr13, which contains the *MYC* gene that is commonly amplified in human lung adenocarcinoma, has 2-copy gains in both cell lines. Both cell lines have 2-copy loss of the most commonly deleted genome segment in human lung adenocarcinoma—which contains the genes *CDKN2A/B/B-AS1* (one focal, the other larger CNA; Fig 1B). BACA has large CNA deletions of the lung adenocarcinoma tumor suppressor *PIK3R1* (<1% of human cases) [36] and the pan cancer tumor suppressor *CASP3*. CLAC has a large CNA gain including *NRAS*, a known lung adenocarcinoma driver in 0.4% of human cases [36]. CLAC also has gain of the lung adenocarcinoma oncogene *CCND1* associated with 4% of human cases.

Because the large CNAs contain very high numbers of presumptive bystander genes, it is not straight forward to evaluate all as potential oncogenic drivers. However, it is possible to study focal alterations—here arbitrarily defined as <3Mb—to implicate known cancer genes and pathways. Table 3 shows the results of Cancer Gene annotation system for Cancer Genomics (CaGe) analysis of all identified CNAs [37]. The table was also annotated by manual analysis to consider whether a gene is likely to be a tumor suppressor or oncogene [38]. Using our conservative criteria to minimize false positives, the total number of genes affected by focal alterations is 263. Of those, 129 genes were called as either cancer drivers (89 genes) or pathway genes (40). We determined that at least 28 and 13 of those genes, respectively, are gained or lost in the predicted direction to be oncogenic (i.e., gain of oncogenes and loss of tumor suppressors). Two of those genes–*CDKN2A* (Fig 1B) and *LRP1B* –are reported to be significantly mutated in lung adenocarcinoma (but the latter, called only in the earlier study, may have been due to the large size of the gene) [36]. Gene Set Enrichment Analysis (GSEA) of all focal CNA genes yielded the top match by significance as genes altered in a complex therapeutic model (GSEA gene set name Zhang_antiviral_response_to_ribavirin_dn) applied to the human lung adenocarcinoma A549 cell line; another top hit was genes down-regulated by stable expression of *SEMA3B* in the human lung adenocarcinoma cell line H1299 (*p*-values/false detection rate *q*-values of 4.5E-06/3.27E-02 and 2.99E-05/4.65E-02). While those mechanistic studies with human cell lines may not be directly relevant to the canine BACA and CLAC cell lines, this finding and the implicated driver genes mentioned indicate that the dog lines are relevant to human lung adenocarcinoma.

Cancer Gene Census (CGC), Wellcome Trust Sanger Institute (http://cancer.sanger.ac.uk/census/; Futreal *et al.* PMID: 14993899); Cancer Gene Index (CGI), National Cancer Institute (https://wiki.nci.nih.gov/display/cageneindex/Cancer+Gene+Index+End+User+Documentation); bold, genes in the correct direction of gain/loss to be oncogenic or tumor suppressive (according to analysis of pan cancer sequence mutation profile by Davoli *et al.* PMID: 24183448); italics and underline, genes that can be either oncogenic or tumor suppressive

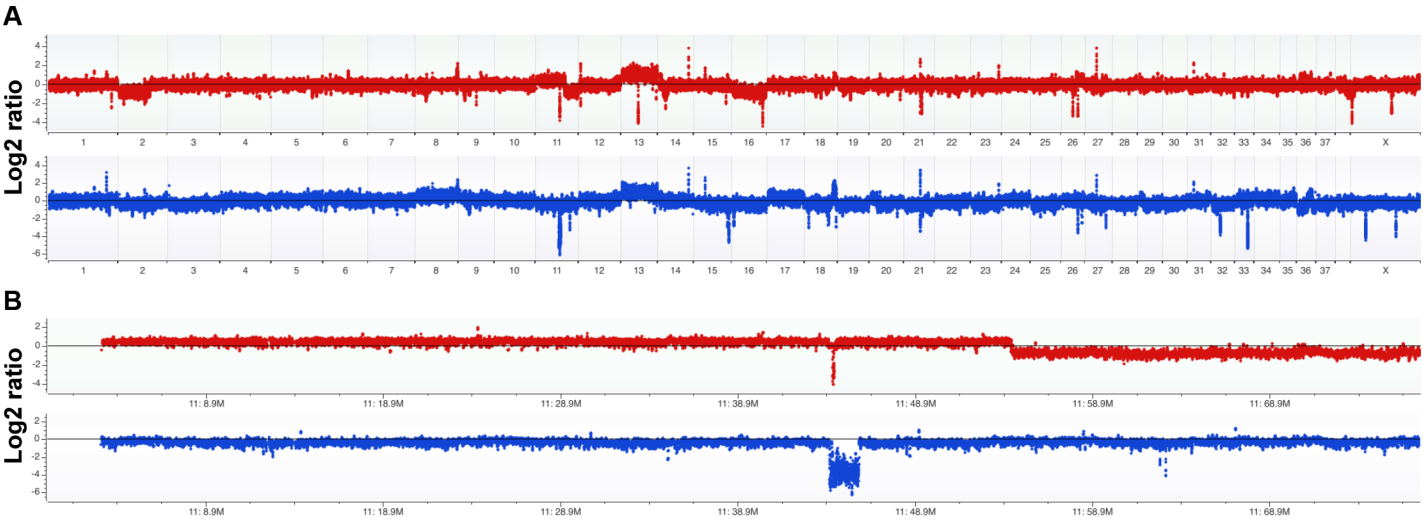


Fig 1. Comparative genomic hybridization analysis of canine BACA and CLAC lung adenocarcinoma cell lines. A custom 1M feature array was used to measure the DNA copy number across the genome relative to germline DNA from a normal dog. The zero line corresponds to a ratio of 1:1 between test DNA and a male reference DNA, or generally a copy number of 2. Gains are above the line and losses below. (A) Whole genome copy number analysis is shown with BACA on top (red) and CLAC on bottom (blue). Among the important findings, both cell lines have a 2-copy loss of *CDKN2A/B/B-AS1* (encoding tumor suppressors p14-p19/ARF and p16/INK4; also shown in close-up in B panels) and a 2-copy gain of chr13 (which contains the oncogene *MYC*). Other large alterations are a 1-copy loss of most of chr2 and chr16 in BACA and a 1-copy gain and loss of chr17 and chr11, respectively in CLAC. Most of all other variants shared by the two cell lines represent Copy Number Variation (CNV) in the reference DNA. (B) Copy number analysis of complete chr11 shows BACA has a 1-copy gain over 2/3rds of its length and a 1-copy loss over the remainder and CLAC has a 1-copy loss over its full length. Both lines have a relatively small 2-copy loss (focal in BACA) that overlaps only *CDKN2A/B/B-AS1* between the two cell lines.

doi:10.1371/journal.pone.0142007.g001

(according to Davoli *et al.* PMID: 24183448); asterisk, genes among 26 significantly mutated genes in human lung adenocarcinoma (FDR < 0.1; according to Ding *et al.* PMID: 18948947)

cDNA transcripts are present for NKX2-1, HSP70 and HSP90 isoforms, and HSP90 client proteins

Reverse transcriptase-polymerase chain reaction (RT-PCR) was used to identify the presence of mRNA for NKX2-1, also known as TTF-1, a marker of lung differentiation, and eleven

Table 3. Gene annotation summary of focal copy number alterations in BACA and CLAC.

Annotation Category	*Genes*
CGI Genes	*TSC2*, **CDK6**, **CDKN2A***, **SMO**, CHEK2, *BCL6*, **LPP**, **NF2**, RANBP17, CRTC3
CGI (not CGC)	**ARHGDIG**, CARD8, **CTSC**, DUSP10, IRF5, MAGI2, **MAPK8IP3**, NRF1, PEX1, **RAB11FIP3**, RPS27L, TUSC3, UBE2E3, UBE2H, UBE2I, AHCYL2, CYP51A1, **KAT2B**, MAD2L1, TNFRSF1B, AXIN1, NME3, PARK2, HMGCS2, SMS, ATP6V1F, **IL1RAP**, FLNC, **SDC1**, SGOL1, CALU, RAB5A, **APOB**, **SSTR5**, CACNA1H, IMPDH1, IQGAP1, **CLDN16**, LEP, CCL20, **CDKN2B**, MPG, NTHL1, CLDN1, DMD, RHOB, AP1B1, **TRIO**, **B3GALNT1**, **GALNT3**, *ARHGAP27*, ATAD3B, BAIAP3, CCDC136, **DNAH5**, DYNLRB2, FAM190A, **GFER**, KRIT1, **LRP1B***, LRRC4, MRPL28, MSLN, MSR1, **NLRP5**, OCA2, OSR1, **PKD1**, PKP4, PNN, RASL10A, RHBDD3, RHBDF1, **SCN9A**, SMC4, **SOX8**, SSU72, **TP63**, **ZFHX3**
CGC & CGI Pathway	ARHGAP15, ARHGAP24, **DECR2**, DIAPH2, GDF7, GNPTG, HERC2, LAPTM4A, NEFH, **PIWIL2**, SEC23A, SLC9A3R2, SPHKAP, STUB1, IGFALS, **PPM1L**, **PIGQ**, **IL12A**, RPL38, RPL3L, GEMIN2, **MASP1**, **GNG13**, OPN1SW, **PPP3CC**, POLR3D, CTNNA3, NDUFB10
CGI (not CGC) Pathway	SLC39A14, HAGHL, SLC19A3, KYNU, **OR14C36**, **OR14I1**, **OR2G6**, OR2T11, **OR2T27**, **OR2T6**, OR51F1, SLC39A14

doi:10.1371/journal.pone.0142007.t003

receptor and cytoplasmic tyrosine and serine/threonine kinases, in the BACA and CLAC canine lung cancer lines. Both cell lines expressed NKX2-1 (Fig 2A). Ten of eleven HSP90 client kinase cDNA transcripts investigated for expression were present in both cell lines. EGFR, c-Kit, HER2, VEGFR, PDGFRα/β, c-MET, MAPK, c-RET, AKT-2 were present and transcripts for ALK were absent (Fig 2A). Both cell lines had transcripts of HSP70 and the HSP90 isoforms (Fig 2B).

STA-1474 inhibits BACA and CLAC cell line proliferation

Cell viability was assessed after 72 h of treatment with increasing concentrations of drugs commonly used to treat lung cancer in humans and dogs (Fig 3A and 3B). IC_{50} values were determined for each drug in both cell lines. Effects on cell proliferation were generally dose- and cell line-dependent. With respect to currently used cytotoxic chemotherapeutics, treatment of the BACA line with vinorelbine achieved the lowest IC_{50} (0.729 μM). In contrast, when the CLAC line was treated with increasing concentrations of vinorelbine for 72 h, an IC_{50} was never reached, i.e. this cell line was relatively resistant (Fig 3A). Treatment of the CLAC line with carboplatin did not result in a dose-dependent decrease in cell viability until a concentration of 100 μM was reached. Increasing doses of carboplatin did result in decreasing BACA cell viability. Gemcitabine ICs_{50} for the BACA and CLAC lines were lower than those achieved with carboplatin. Statistically significant decreases in BACA cell viability at 72 h were present for all concentrations of gemcitabine used. The IC_{50} for gemcitabine-treated CLAC cells was approximately three times greater than that of BACA. Overall, BACA cell viability was more sensitive to these cytotoxic drugs as indicated by the lower ICs_{50} compared to the CLAC line.

Of the three different small molecule inhibitors evaluated, STA-1474, the HSP90 inhibitor, achieved the greatest inhibition of cell viability and had the lowest ICs_{50} after 72 h of drug exposure (Fig 3B). BACA cell viability was minimally affected by torceranib phosphate treatment. A concentration-dependent decrease in CLAC viability was found with increasing concentrations of toceranib resulting in a lower IC_{50} (0.47μM) when compared to BACA. Crizotinib treatment of the CLAC cell line did not produce a statistically significant decrease in viability which resulted in an unachievable IC_{50}. Although critzotinib treatment of the BACA cell line produced a significant ($^{***}p<0.001$) decrease in viability at both the 1.4 μM and 1.6 μM concentrations, the cell viability was higher than 50% with the doses used.

STA-1474 promotes apoptosis in a time- and dose-dependent manner

To determine if the growth inhibitory effect of STA-1474 on both cell lines was associated with apoptosis, cell lines were treated with increasing concentrations of STA-1474 (0.005–0.05 μM) for 24 h and evaluated for annexin V and propidium iodide (PI) staining. Although incubation of the BACA line with 0.05 μM STA-1474 for 24 h increased the proportion of cells undergoing apoptosis (represented by both annexin V and annexin V & PI positivity), this was not statistically significant (Fig 4A). In contrast, 24 h exposure of the CLAC line with 0.05 μM STA-1474 resulted in a significant ($p<0.05$) increase in the proportion of apoptotic cells. Characterization of the identified apoptotic response was evaluated further by quantifying the executioner caspase activity of caspases 3 and 7 after 24 and 48 h exposures to vehicle (DMSO) or increasing concentrations of STA-1474 (0.05–1 μM). Twenty-four and 48 h treatments of the BACA line with STA-1474 resulted in a significant dose-dependent increase in caspase activity (Fig 4B, middle left panels). A dose-dependent significant increase in caspase 3/7 activity was seen in the CLAC line after 48 h of drug exposure (Fig 4B, middle right panels). The relative increase in caspase3/7 increased significantly ($p<0.001$) between 24 and 48 h in the CLAC line but not in the BACA line (Fig 4B, lower panels).

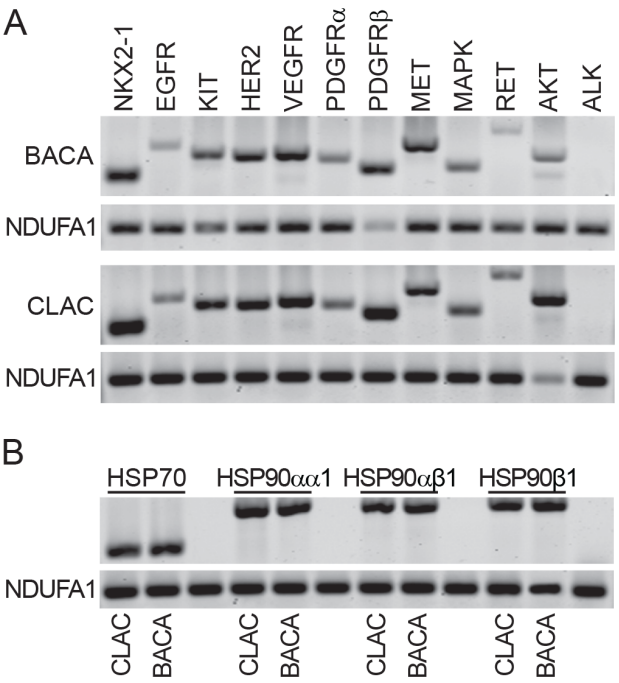


Fig 2. Characterization of BACA and CLAC canine lung cancer lines by RT-PCR. (A) Reverse transcriptase cDNA transcripts of NKX2-1 (lung cancer cell marker) and HSP90 client proteins in BACA and CLAC cell lines. NDUFA1 serves as a loading control. (B) Reverse transcriptase cDNA transcripts of HSP70 and three HSP90 isoforms in BACA and CLAC cell lines. NDUFA1 serves as a loading control.

doi:10.1371/journal.pone.0142007.g002

STA-1474 decreases expression of signal transduction proteins and up-regulates HSP70 expression in canine lung cancer cells in a dose-dependent manner

To further characterize the *in vitro* activity of STA-1474 in comparison to torceranib phosphate, we assessed the ability of these compounds to deplete critical client proteins of HSP90 and multiple receptor tyrosine kinase targets of torceranib. We also evaluated the ability of these compounds to extinguish downstream signaling of the PI3K/mTOR/S6 and RAF/MEK/ERK pathways and the transcription factor, STAT3. Treatment of the BACA cell line with biologically relevant concentrations of STA-1474 induced client-protein depletion of HER2 (starting at 0.05 μmol/L) and STAT3 (at 1.0 μM). Degradation of the phosphorylated forms of the downstream signaling proteins, pAKT, pMAPK, pS6, and transcription factor, pSTAT3, occurred after exposure to 1.0 μM in the BACA cell line (Fig 5A); whereas in the CLAC cell line, decreased phosphorylated forms of the same proteins occurred at just 0.5 μM (Fig 5B). Interestingly, protein levels of pmTOR were unaffected by exposure to biologically achievable doses of STA-1471 in the BACA line and only slightly affected the CLAC line. Toceranib treatment of the BACA cell line resulted in a slight decrease in HER2. Degradation of all other proteins and phosphorylated forms were unaffected at exposures <10 μmol/L (Fig 5A and 5B).

As inhibition of HSP90 typically leads to the increased expression of other HSP family members that can be used as surrogates for HSP90 inhibition, we evaluated the effect of HSP90 abrogation on HSP70 expression. Treatment of both cell lines with increasing concentrations of STA-1474 (0.005–1.0 μM) for 48 h resulted in an increase in HSP70 expression at 0.5 and 1.0 μM exposures (Fig 5C, left panels). As expected, treatment of both cell lines with increasing concentration of toceranib did not increase HSP70 expression (Fig 5C, right panels).

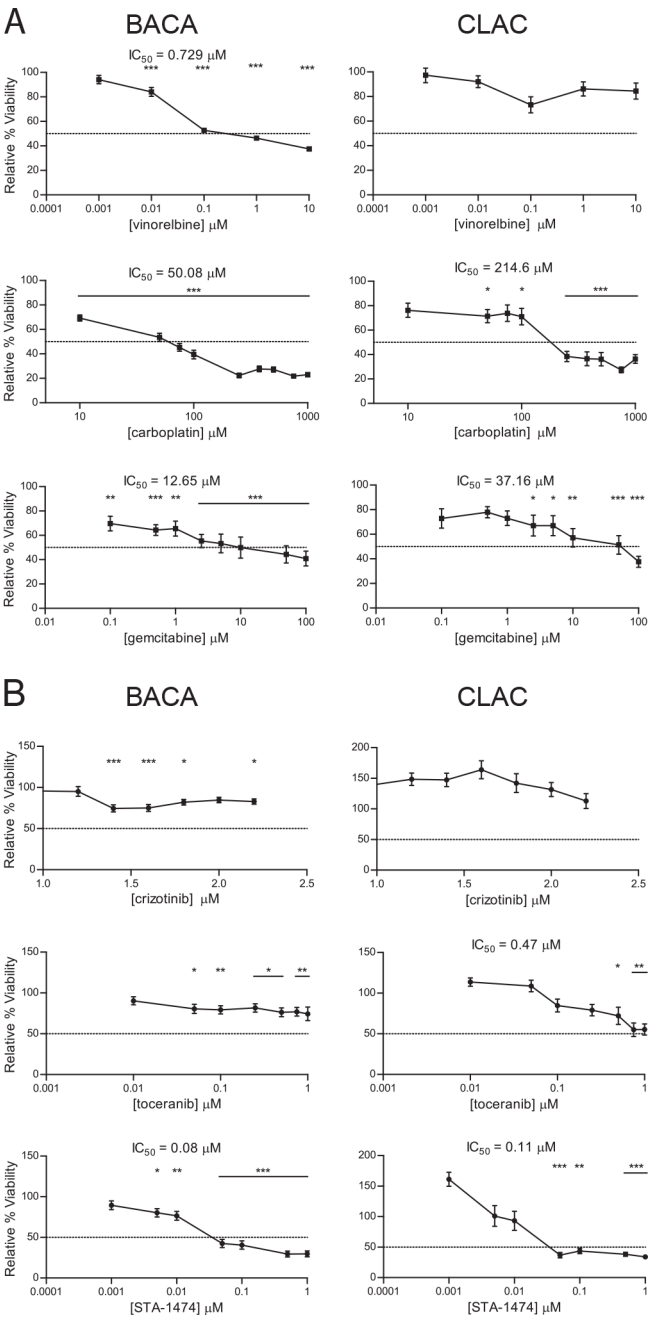


Fig 3. Viability assays and half-maximal inhibitory concentrations (IC_{50}) for canine lung cancer cell lines. Cells were treated with increasing concentrations cytotoxic drugs (A) or small molecules inhibitors (B) and proliferation was evaluated after 72 h. Treatment effects were normalized to the drug vehicle-treated control group. Each graph shows mean±SEM. IC_{50} were calculated for each experiment. The dotted line in the y-axis represents the 50% relative viability. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

doi:10.1371/journal.pone.0142007.g003

Synergistic cytotoxicity from dual inhibition of HSP70 and HSP90 is dose and cell-line dependent

The binding of a client protein to HSP90 requires the cooperation of HSP90 with other chaperone proteins, HSP70 and HSP40. As induction of HSP70 was present in both cell lines after 72

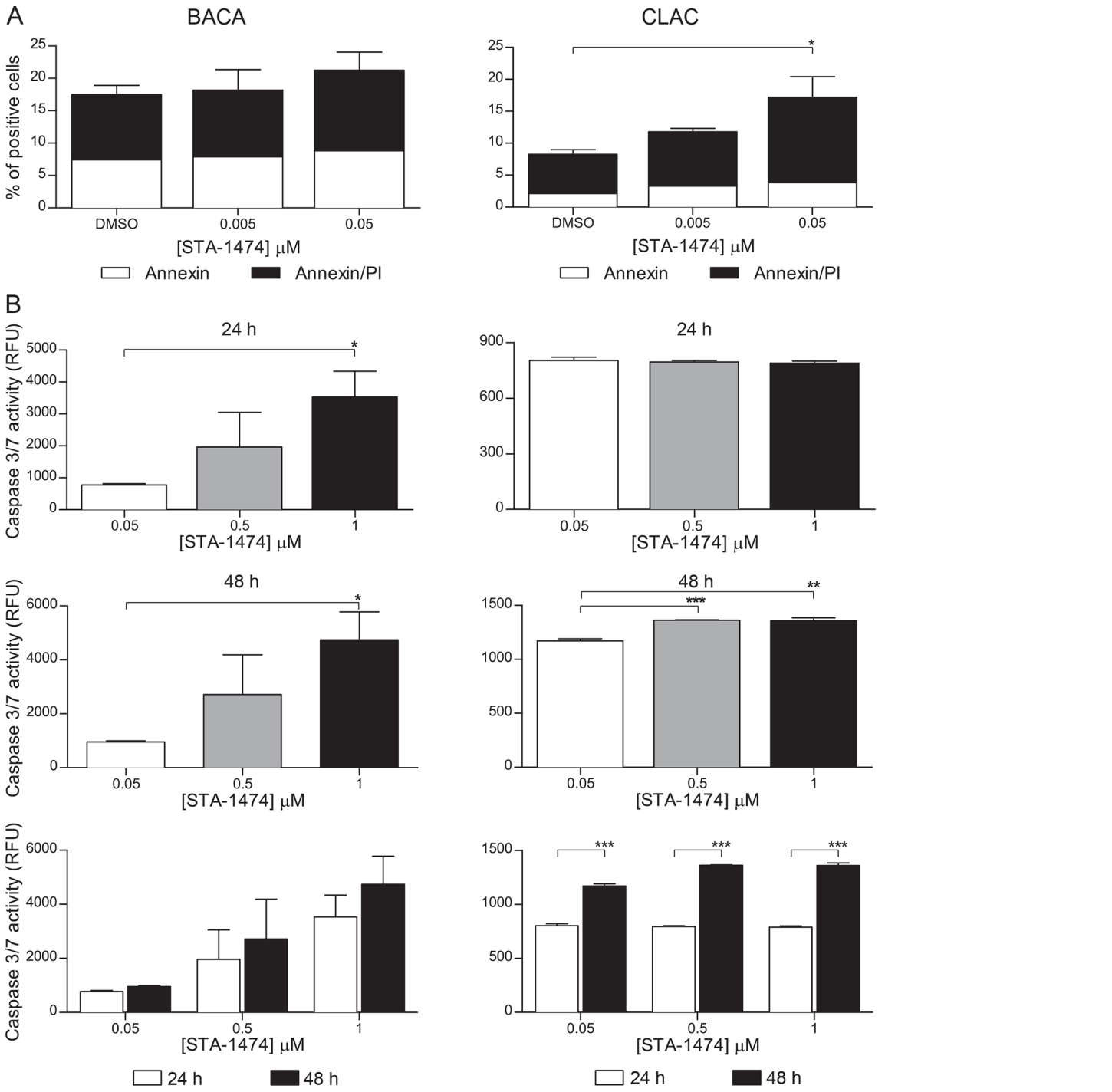


Fig 4. Evaluation of apoptosis in canine lung cancer cell lines treated with STA-1474. Apoptosis was assessed by annexin V/PI staining flow cytometry and detection of caspase 3/7 enzymatic activity. (A) Cells were treated with dimethyl sulfoxyide (DMSO, control) or 0.005–1 μM of STA-1474 for 24 h. Staining with annexin V and the vital dye, propidium iodide (PI), were used to evaluate early and late apoptosis. Cells that are considered viable are both annexin V and PI negative, while cells that are in early apoptosis are annexin V positive and PI negative, and cells that are in late apoptosis or already dead are both annexin V and PI positive. (B) Both cell lines were treated as above and evaluated for executioner caspase-mediated apoptosis. Activated caspases 3 and 7 were assessed 24 and 48 h after treatment. Experiments were performed in triplicate and repeated three times. Each graph shows mean ± SEM. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

doi:10.1371/journal.pone.0142007.g004

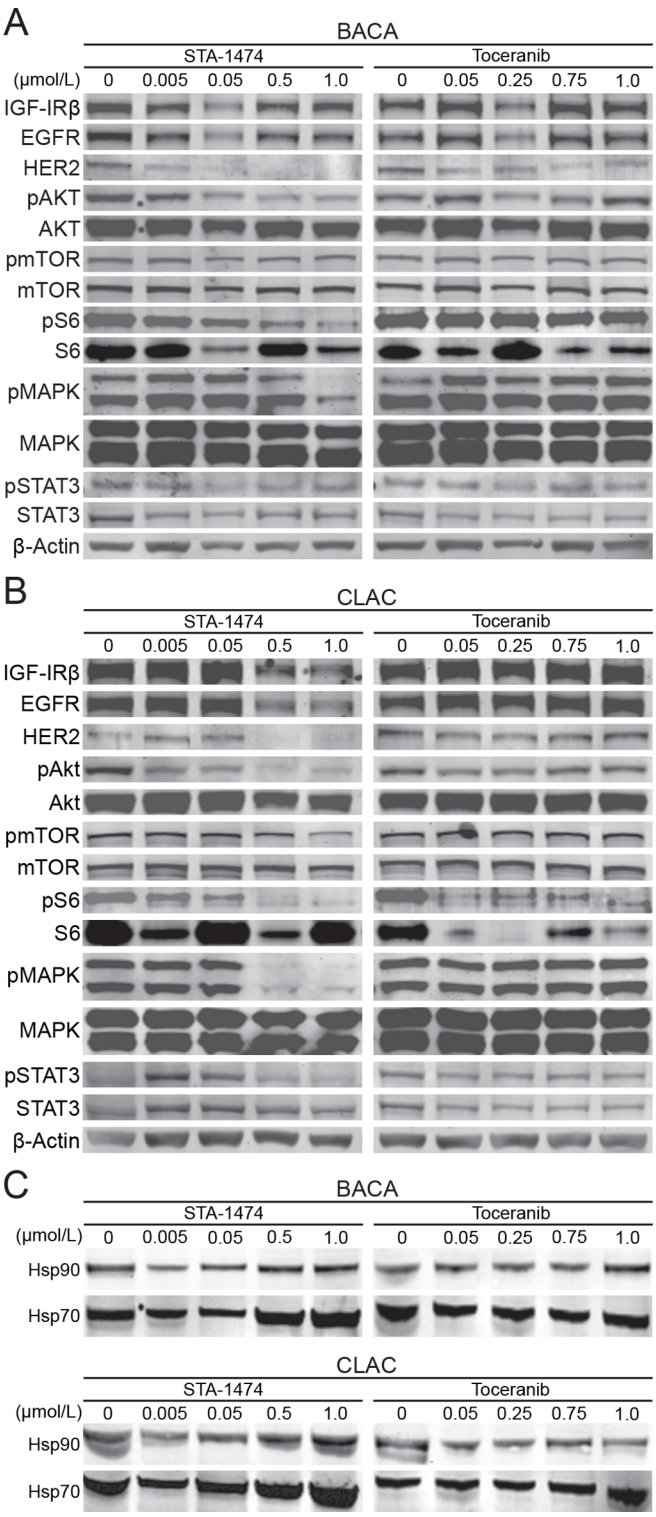


Fig 5. Protein expression of HSP90-regulated proteins, HSP70 and HSP90 in canine lung cancer cell lines after treatment with small molecule inhibitors. A set of three plates for each cell line was used for evaluation of total and phosphoproteins of downstream signaling pathways (A) Representative immunoblots of HSP90 client protein expression from BACA whole cell protein lysates after treatment with STA-1474 or toceranib phosphate. (B) Representative immunoblots of HSP90 client protein expression from CLAC whole cell protein lysates after treatment with STA-1474 or toceranib phosphate. Controls were cell lines treated

with the drug solvent, DMSO, as represented by the 0 concentration. Evaluation of phosphoprotein forms of the proteins are indicated by "p". Drug concentrations are μmol/L. The β-actin Western blots serve as loading controls. (C) Immunobloting from whole cell protein lysates of HSP70 and HSP90 of BACA and CLAC lines treated with DMSO (control), STA-1474 and toceranib phosphate.

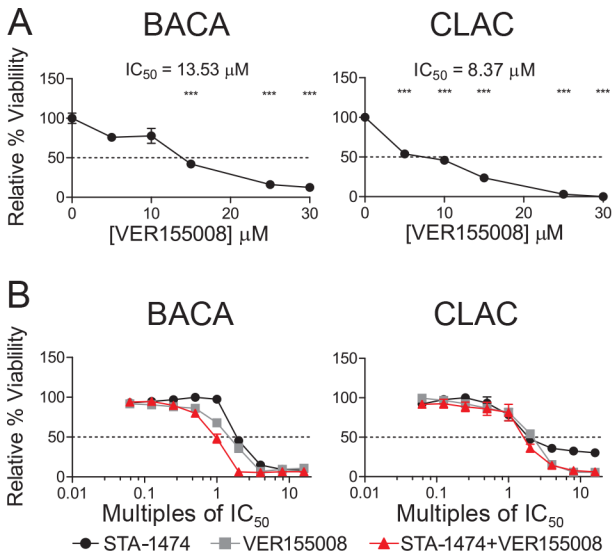
doi:10.1371/journal.pone.0142007.g005

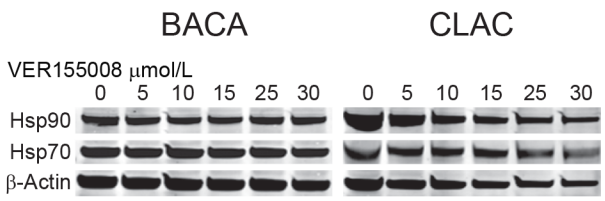
h treatment with STA-1474, we sought to determine if synergistic cytotoxicity would occur when cells were treated with the combination of a HSP90 inhibitor and a HSP70 inhibitor. Cell line viability was assessed after 72 h of treatment with increasing doses of the HSP70 inhibitor, VER155008 (0–30 μM). There was a dose-dependent reduction in viability of both cell lines after treatment with VER155008 without a change in the expression of HSP70 or HSP90 in either cell line (Fig 6A). Treatment of BACA and CLAC cells for 72 h with a combination of VER155008 and STA-1474 inhibitors in concentrations that were constant ratios of multiples of the ICs_{50} reduced relative cell viability in a dose-dependent manner (Fig 6B). Dose-effect and associated CI values for the drug combination treatment (VER155008/STA-1474) at constant ratios are presented in Fig 6C. The model is most accurate at the $f_a = 0.50$, the point that the drugs affect 50% of the cells. BACA cells treated with VER155008/STA-1474 combinations were synergistic at $f_a = 0.50$, and with increasing two-fold IC_{50} concentration combinations, the CI values increased to produce moderate synergism at effect levels $> 50\%$. The dose reduction index (DRI) was calculated and is a measure of how many folds the dose of each drug in a synergistic combination may be reduced at a given effect level when compared with the doses of each drug alone. The BACA line first showed evidence of synergism to the drug combination at $f_a = 0.50$ with a CI of 0.70. The DRI at $f_a = 0.50$ was 1.16 and 3.90 fold for VER155008 and STA-1474, respectively. The VER155008/STA-1474 treatment in the CLAC line provided synergism at the $f_a = 0.50$, and progressed to antagonistic at effect levels $> 90\%$. The DRI at $f_a = 0.50$ was 1.5 and 10.5 fold for VER155008 and STA-1474, respectively. Unfortunately, the VER155008 concentration needed to decrease cell viability was too high to take advantage of the drug combination in terms of obtaining the desirable therapeutic effect by administering low doses of the drugs in combination. These results demonstrate that VER155008/STA-1474 combination has a limited dose range in which synergism is obtained and the concentrations are cell line dependent.

BACA and CLAC cells were plated in DMEM media and were exposed to the compounds for 72 h at constant fixed ratios of [75:1](VER155008:STA-1474 CLAC) and [169:1] (VER155008:STA-1474 BACA) and relative cell viability assessed. The resulting CI values are shown for effect level. For example, the doses of these two drugs needed to achieve a 97% decrease in BACA relative viability gives a CI value that would indicate moderate synergism which provides a drug reduction index score indicative of a four-fold decrease of VER155008 and two-fold decrease of STA-1474.

Relative cell viability after treatment with STA-1474 differs in monolayer cultures *vs*. tumor spheroids and STA-1474 decreases tumor-stromal fibroblast viability

To determine the efficacy of STA-1474 on canine lung cancer cells in different model systems, we evaluated the STA-1474 ICs_{50} for cells grown as monolayers or as three-dimensional tumor spheroids. Both BACA and CLAC cells formed tumor spheroids when grown for three days in ultra-low attachment well plates (Fig 7A). BACA cells formed compact, tight spheroids, whereas CLAC cells formed loose spheroids. Different STA-1474 ICs_{50} were obtained depending on the model used. Monolayer cultures were seeded and then allowed to grow for 72 h before treatment to mimic the growth period needed for the formation of spheroids. BACA





C

VER155008 and STA-1474		
BACA		
Effect Level (%)	CI	Description
50	0.70	Synergism
75	0.73	Moderate Synergism
90	0.78	Moderate Synergism
97	0.77	Moderate Synergism
CLAC		
Effect Level (%)	CI	Description
50	0.69	Synergism
75	1.04	Nearly Additive
90	1.10	Nearly Additive
97	2.05	Antagonism

Fig 6. Relative cell viability assays, half-maximal inhibitory concentrations (IC_{50}s) after treatment with an HSP70 inhibitor (VER155008) or a combination of STA-1474 and VER155008. (A) Cells were treated with HSP70 inhibitor, (VER155008) for 72 h. The relative viability and IC_{50} were determined for both cell lines after treatment. (B) BACA or CLAC cells were plated in DMEM media for 24 h and then treated for 72 h with VER155008 only (gray squares), STA-1474 only (black circles) or a combination of VER155008 and STA-1474 concentrations (red triangles) ranging from 0.0625X to 16X their IC_{50} concentrations. Each graph shows mean ± SEM. Each group was compared to the DMSO control. The dotted line in the y-axis represents the 50% relative viability. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$. (C) Multi-drug combination dose-effect analysis for the doublet combination of VER155008 and STA-1474 on BACA and CLAC cell lines as measured by the combination index (CI). The CI value definitions are represented in the description column.

doi:10.1371/journal.pone.0142007.g006

cells grown as a monolayer, demonstrated a dose-dependent decrease in viability with an IC_{50} of 0.168 μM and <5% of cells viable at drug concentrations of 0.5 μM or higher. When grown as spheres, BACA cells were less sensitive to STA-1474 drug treatment, resulting in a doubling of the IC_{50} (0.348 μM) compared to monolayer cultures. CLAC cells grown as a monolayer had a dose-dependent decrease in viability that was significant at doses of 0.5 μM or higher. CLAC spheroid cultures were resistant to all drug treatment concentrations investigated (up to 10 μM; Fig 7B). We also evaluated the relative cell viability of tumor-stromal fibroblasts after 72 h of treatment with STA-1474. There was a dose-dependent decrease in cell viability that was significant with concentrations of ≥ 0.1 μM. The determined IC_{50} was higher than for both canine lung cancer cells lines (Fig 7C *vs*. Fig 3B).

Discussion

The prognosis for dogs with advanced lung cancer remains poor and new treatment options are needed. The molecular characterization of canine lung cancer is limited compared to human lung cancer. In this study we wanted to investigate in canine lung cancer cells the activity of some of the molecularly targeted therapies used to treat lung cancer in people, as well as investigate other chemotherapy drugs used to treat lung cancer in dogs. The biologic activity of these different chemotherapies and small molecule inhibitors on canine lung cancer cell lines has not been reported previously and could provide guidance and insight into the choice of drugs for treating dogs with lung cancer.

Recent reports of HSP90 inhibitor activity in human non-small cell lung cancer (NSCLC) phase IIb/III clinical trials have provided compelling rationale for investigating the feasibility of using HSP90 inhibitors for treatment of lung cancer in dogs [24]. HSP90 is a highly conserved protein that folds newly synthesized proteins into their biologically active conformations preventing aggregation. HSP90 also maintains cellular protein homeostasis by acting as a

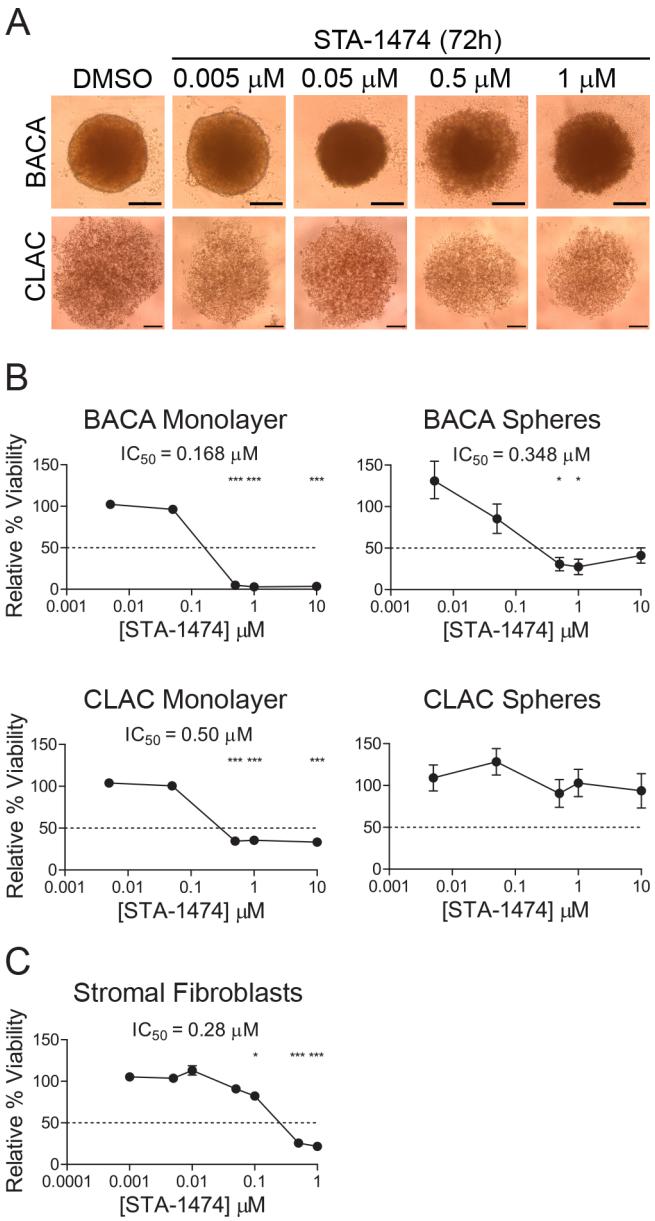


Fig 7. Viability and half-maximal inhibitory concentration (ICs_{50}) of STA-1474 treated canine lung cancer cell-line derived tumor spheroids, monolayers and tumor-stromal fibroblasts. (A) Tumor spheroids derived from both cell lines were allowed to form for 72 h after seeding cells in the ultra-low attachment wells. Images taken of one field of view/well on an inverted microscope at 40X magnification, scale bar represents 200 μM (B) Immediately after formation, tumor spheroids were treated with DMSO or increasing concentrations of STA-1474 for 72 h. Tumor spheroids and monolayers from each cell line were grown and treated identically. Cultured cells were treated with STA-1474 for an additional 72 h and ICs_{50} were calculated. (C) Tumor-stromal fibroblasts were seeded, allowed to form a monolayer for 24 h then incubated with STA-1474 for 72 h and viability determined. Experiments were performed in four replicates and repeated twice. Each graph shows mean ± SEM and each group was compared to DMSO. The dotted line in the y-axis represents the 50% relative viability. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

doi:10.1371/journal.pone.0142007.g007

molecular chaperone with its action modulated by co-chaperones and client proteins [15–17]. As HSP90 regulates multiple signaling cascades, the effects of pharmacological blockade of HSP90 should interfere with a variety of client proteins and biochemical pathways. Given that we know very little about the current crucial signaling pathways for canine lung tumor viability and considering we could target multiple signaling proteins with HSP90 inhibition, this prompted us to evaluate the pro-drug, STA-1474, in these canine lung cancer cell lines. Our initial experiments were designed to characterize the suitability of our cell lines to serve as appropriate *in vitro* models that would potentially respond to therapeutic intervention by inhibition of selected targets.

Somatic mutations are the predominant mechanism that gives rise to cancer. The average cancer cell has approximately four sequence mutations of oncogenes (mean 1) and tumor suppressors (3), 11 very large CNAs involving whole chromosomes (2 gain, 2 loss) or chromosome arms (3 gain, 5 loss), and 23 focal CNAs (11 gains, 12 losses)[38]. Here we have determined the CNAs for BACA and CLAC using high resolution array CGH. The most striking findings were 2-copy gains of chr13, which contains *MYC*, and 2-copy loss of a small part of chr11, for which the overlapping segment between the two cell lines includes only *CDKN2A/B/B-AS1*. These are among the most common CNAs seen in human lung adenocarcinoma [36], with *CDKN2A* being involved in 43% of cases. Notably *CDKN2A* encodes ARF, which directly interacts with overexpressed MYC protein to block its transformation and proliferation activities [39]. Other large CNA genes gained or lost in the correct direction to drive cancer are *PIK3R1* and *CASP3* in BACA, and *NRAS* and *CCND1* in CLAC [36]. Table 2 shows the many cancer driver and pathway genes affected in the cell lines by focal CNAs, including *TSC2*, *NF2*, *BCL6*, *CHEK2*, *CDK6*, *KAT2B*, *PKD1* and *TP63*. Among the mechanistic insights, the CNAs suggest relevance for HSP90 inhibition that would be expected to have therapeutic effects through the PI3K and MAPK pathways (e.g., *PIK3R1*, *TSC2*, *BCL6*, *NF2*, *PKD1*, and *NRAS*). Other pathways that are likely to be affected according to our findings are cell cycle progression (e.g., *CDKN2A* and *CCND1*) and apoptosis (*CASP3*). The array CGH data thus support both lung adenocarcinoma and pan cancer relevance. Additionally, the GSEA analysis of these two canine cell lines strongly implicates human lung adenocarcinoma among all other cancer data. These facts formally demonstrate that BACA and CLAC can serve as comparative oncogenomic models for development of drug treatments to mammalian lung adenocarcinomas in which a set of common driver mechanisms are present.

Once we confirmed that the cell lines were appropriate *in vitro models* to evaluate candidate drugs that have been effective in controlling human NSCLC, it was important to establish that these canine lung cancer cell lines expressed HSP90 isoforms, HSP70, and various receptor tyrosine kinases and downstream kinases, relevant to lung cancer in humans. These included EGFR, c-Kit, HER2, VEGFR2, PDGFRα, PDGFRβ, c-Met, MAPK, c-Ret and Akt. The cell lines used for this work were from primary canine lung tumors and of pulmonary cell differentiation as the presence of cDNA transcripts for TTF-1, a nuclear protein [40]. TTF-1, also known as NKX2-1, is a homebox-containing transcription factor essential for the development of the lung, and its use as a marker of lung adenocarcinoma, has been recommended by the newer classifications of human NSCLC expressed in follicular cells of the thyroid gland and pneumocytes was present. TTF-1 has a specificity of 100% and sensitivity of 85% in canine primary lung cancer [10, 11, 41]. Both cell lines had cDNA transcripts for HSP70, HSP90, and all the client proteins and other kinases investigated except ALK.

When complete surgical resection of a primary lung tumor is not possible in canine patients, treatment with cytotoxic chemotherapy may be considered in an attempted to slow the progression of the disease. Interestingly, the cytotoxic drugs used in this study had limited to no apparent effect on cell viability *in vitro*, or only had effects at drug concentrations not thought

to be biologically relevant, or achievable, *in vivo*. For example, the IC_{50} of vinorelbine for BACA was 0.72 μM, which is 4- to 10-fold higher than in human NSCLC [42] and the CLAC cell line was drug-resistant. The responsiveness of our cell lines is consistent with what has been reported in the clinical setting, where only two out of seven dogs with macroscopic bronchoalveolar carcinoma had a partial response to vinorelbine treatment [43]. The obtained CLAC IC_{50} dose for carboplatin is not tolerable in dogs. The tolerated carboplatin IC_{50} has been extrapolated from pharmacokinetic data performed in laboratory beagle dogs using C_{max} at the recommended dose of 300 mg/m^2 given as an IV bolus [44]. An estimated maximum tolerated dose of carboplatin is 250 μM [44]. Likewise, a previous pharmacokinetic study evaluating gemcitabine in dogs found the maximum tolerated dose of 22 mg/kg resulted in a C_{max} of 20–30 μg/mL which is equivalent to a molarity of 67 to 100 μM [45]. When the carboplatin ICs_{50} for the lung cancer cell lines (50 and 214 μM for BACA and CLAC, respectively) is compared to ICs_{50} for other cancer cell lines of canine origin, they are higher than then those for canine mammary gland tumors or canine melanoma (30.5 μM and 6.1 μM, respectively) [46, 47]. The gemcitabine IC_{50} differed between the cell lines, with the CLAC IC_{50} being three-fold higher than BACA IC_{50}. Our IC_{50} findings for gemcitabine are similar to the ICs_{50} reported for canine osteosarcoma cell lines, which ranged from 5.7 to 15.3 μM for the cell lines that had dose-dependent decreases in cell proliferation [48].

We evaluated the prodrug STA-1474 in lieu of ganetespib because of its greater solubility in water which facilitates its use in the dog without untoward side effects. A phase I study evaluated STA-1474 in dogs with solid tumors and reported measurable objective responses for malignant mast cell disease, osteosarcoma, melanoma and thyroid carcinoma [26]. Based on the favorable evidence of STA-1474 displaying potent activity against both canine lung cancer cell lines, we further investigated the effects of HSP90 inhibition in canine lung cancer. STA-1474 decreased cell viability at 72 h, induced apoptosis and promoted activation of caspase 3/7 in a dose- and time-dependent manner in both cell lines. Apoptosis was detected 24 h after treatment and caspase 3/7 activity continued to increase during the first 48 h. As many of the HSP90 client proteins are needed for cell survival and proliferation [49] use of HSP90 inhibitors induces both cell cycle arrest and apoptosis in cancer cells [50]. Similar results have been seen in canine osteosarcoma and mast cell tumor cell lines treated with either ganetespib or STA-1474, inducing growth inhibition that was at least in part mediated by caspase 3/7-dependent apoptosis [51, 52].

Ganetespib, a second generation HSP90 inhibitor, effectively and simultaneously destabilizes HSP90 client proteins in NSCLC cells including receptor tyrosine kinases and canonical JAK/STAT, PI3K/AKT, MAPK and mTOR signaling. Moreover, ganetespib accumulates in tumors relative to normal tissues, with a half-life in the tumor that is 10- to 19-fold longer than in normal tissues or plasma [23]. Our experiments show that the prodrug of ganetespib, STA-1474, down-regulates signal transduction proteins in a dose-dependent manner. For some proteins (AKT, MAPK, STAT3) only the phosphorylated form is decreased but not the total protein. Lack of total STAT3 down-regulation has been reported before with STA-1474 treatment [51], which is an unexpected finding as STAT3 is a HSP90 client protein. This result can be explained when individual protein turnover rate is considered. HSP90 inhibition should lead to the rapid degradation of newly synthesized proteins and those with a longer half-life will show a slower decrease in loss of protein after the HSP90 inhibition [53].

The tyrosine kinase inhibitor toceranib is a multi-targeted kinase inhibitor that effects both tumor cell proliferation and tumor angiogenesis. We sought to determine if common tyrosine kinases which support lung tumor growth could be inhibited by torceranib. Even after 72 h of exposure to the highest dose of torceranib used (1 μM) this compound was unable to abrogate the phosphorylated and total protein expression of all the kinases evaluated in both cell lines.

Further studies are needed to evaluate if torceranib demonstrates inhibitory activity against VEGFR, PDGFR and c-Kit protein expression in canine lung cancer cell lines.

Exposure of tumor cells to HSP90 inhibitors induces a cellular protective and compensatory response which is to increase the expression of other heat shock proteins, notably HSP70. The increase of HSP70 expression has been shown to abrogate the extent of cell death [50]. Indeed we saw a reciprocal upregulation of HSP70 when both cell lines were treated with increasing doses of HSP90 inhibitors. Previous studies have reported an increase in HSP70 protein expression when canine tumors and cancer cell lines were treated with STA-1474 [26, 51].

HSP70 has also been shown to be a driver of oncogenesis therefore therapy using a combination of a HSP90 inhibitor with a HSP70 inhibitor may provide a wider therapeutic window and increase the target-driven therapeutic index. We wanted to investigate if the simultaneous inhibition of HSP70 would enhance the growth inhibitory effects of HSP90 treatment. The use of VER155008, a HSP70 inhibitor, suppressed cell growth in both cell lines but only at high μM concentrations. Similar results have been described in human NSCLC cells treated with this drug [54]. Surprisingly when the cell lines were treated with the fixed ratio combinations of STA-1474 and VER155008 based on their IC_{s50}, the doublet combination effects were different not only between cell lines but also in terms of the desired effect level. Moderate synergy was seen when the BACA line was treated with the combination at concentrations higher than its' IC_{50} values, whereas in the CLAC line this combination created synergism only at the IC_{50} value. This was an unexpected finding, as VER155008 has shown to be synergistic with other HSP90 inhibitors in human NSCLC lines [54–56]. Multiple drugs can compete with each other for the same transporter, molecular target or have conflicting effects on the cell cycle. Studies suggest that cancer cells are sensitive to multiple drugs at a certain drug mixing ratio, and that the optimal mixing ratio must be retained in tumor tissues to achieve the maximal drug combination effect [57–59]. The lack of a constant synergistic effect at all levels with the combination therapy may be due to the induction of other protective HSPs or upregulation of alternative oncogenic compensatory pathways not examined in this study.

Evaluation of drug activity using a two-dimensional monolayer of tumor cells poorly models the disease *in vivo*. Tumors are three-dimensional (3D) complex tissues composed of neoplastic cells, vasculature and tumor stroma. As such, we sought to determine the response of STA-1474 in a 3D-model of canine lung cancer as well as in cells of the tumor stroma, represented by tumor-stromal fibroblasts obtained from a primary canine lung tumor cell culture. Both cell lines grew as spheres in ultra-low attachment wells, as described by Vinci *et al.* [60]. In contrast to other studies, the need to add growth factors to the media to foster the formation of spheroids was not needed [61]. This was critical for the interpretation of the ICs_{50} as addition of growth factors can enhance cell proliferation and viability making direct comparison of spheroid and monolayer ICs_{50} impossible. Spheroid conformation was different between the cell lines. This finding was similar to previous reports that have reported different cell lines from a variety of tumors grow in different spheroid patterns [60, 62]. The spheroid shape may influence drug responsiveness. Morphometric analysis of sphere area would be helpful to determine relative cell viability after drug exposure. However, although the sphere area can be measured using image analysis software, determination of area can be ambiguous in approximately 50% of the evaluated cases [62]. The STA-1474 ICs_{50} were higher for both spheroid cell cultures than for the monolayer adherent cultures. This was not unexpected as in general chemotherapy drugs have increased potency in 2D models compared with spheroid cultures, although exceptions do exist with a better response seen in 3D models to some pharmaceuticals [62]. The drug effectiveness in limiting spheroid viability is in part due to the ability of the agent to diffuse into the sphere. Therefore, although spheroids are a better model to evaluate cellular responsiveness to drugs, caution must be used when interpreting the ICs_{50} as lack of

vascularization responsible for the delivery and distribution of the agent throughout the sphere is not present.

The tumor stroma is composed of cancer-associated fibroblasts (CAFs), cells that are phenotypically and functionally different from their normal counterparts. CAFs can induce cancer cell stemness as well as epithelial to mesenchymal transition to promote tumor progression [63]. CAFs can also induce therapeutic resistance in NSCLC [64]. Moreover, cytotoxic treatment can increase both CAF percentage and cytokine secretion, especially IL-17A, both of which contribute to cancer-initiating cells growth and therapeutic resistance [65]. However, stromal fibroblasts also respond to the neoplastic epithelial cells by expressing growth factors [66]. Because FGFR3 is a known HSP90 client protein [67] we treated tumor stromal fibroblasts with STA-1474. Although the IC_{50} for tumor stromal fibroblasts was higher than the cell lines it was still within a biologically achievable dose. This result was dissimilar to what has been reported with other HSP90 inhibitors, which have decreased CAFs cytokine production but not viability [68].

Conclusions

According to our genetic and biochemical analyses, canine lung adenocarcinoma cell lines are relevant to the same human cancer. Likewise, the results of the canine aCGH and the variability in biological response to the chemotherpy agents found in these two cell lines suggests that the canine patient would benefit from precision medicine which has significantly improved the quality of life of the human lung cancer patient. We showed that treatment with STA-1474 decreases BACA and CLAC viability and HSP90 client protein expression at biologically relevant doses, by inducing dose- and time-dependent apoptosis. In contrast with other small molecules inhibitors, STA-1474 affects several proteins in different cellular signaling pathways and decreases viability not only in tumor cells, but also stromal fibroblasts. The efficacy of two small molecule inhibitors in canine lung cancer cell lines grown in a 3D spheroid format established that ICs_{50} will be increased when compared to the ICs_{50} obtained when cells are grown in a monolayer. This finding re-enforces the importance of evaluating cellular responses to drugs in a model that more accurately mimics the natural tumor environment. Nevertheless, the preclinical activity profile of STA-1474 demonstrated in this study provides preliminary evidence that this compound is superior to most of the currently available drugs and may offer an effective therapeutic opportunity to manage the canine lung cancer patient.

Supporting Information

S1 Table. Comparative genomic hybridization analysis. Contains segmentation, classification of focal *vs*. large alterations and gene annotation.
(XLSX)

Acknowledgments

The authors would like to thank Dr. Joelle Finger and Ms. Pepa Domenech-Cabrera for help with flow cytometry experiments and fibroblast viability assays, respectively. Also the authors wish to thank Dr. David A. Proia of Synta Pharmaceuticals Corporation for kindly providing STA-1474, Dr. Joseph J. Wakshlag of Cornell College of Veterinary Medicine for the BACA cell line and Mr. Marc Hardman for help with the figures.

Author Contributions

Conceived and designed the experiments: GL WCK CEA JLR. Performed the experiments: FCV CEA JLR SR. Analyzed the data: GL FCV CEA JLR. Contributed reagents/materials/analysis tools: CAL CEA WCK JLR. Wrote the paper: GL FCV WCK CEA. Obtained permission for use of the BACA cell line: WCK.

References

1. Fleming JM, Creevy KE, Promislow DE. Mortality in north american dogs from 1984 to 2004: an investigation into age-, size-, and breed-related causes of death. Journal of veterinary internal medicine / American College of Veterinary Internal Medicine. 2011; 25(2):187–98. doi: 10.1111/j.1939-1676.2011.0695.x PMID: 21352376.
2. O'Neill DG, Church DB, McGreevy PD, Thomson PC, Brodbelt DC. Longevity and mortality of owned dogs in England. Veterinary journal. 2013; 198(3):638–43. doi: 10.1016/j.tvjl.2013.09.020 PMID: 24206631.
3. Adams VJ, Evans KM, Sampson J, Wood JL. Methods and mortality results of a health survey of purebred dogs in the UK. J Small Anim Pract. 2010; 51(10):512–24. doi: 10.1111/j.1748-5827.2010.00974.x PMID: 21029096.
4. Dobson JM, Samuel S, Milstein H, Rogers K, Wood JL. Canine neoplasia in the UK: estimates of incidence rates from a population of insured dogs. J Small Anim Pract. 2002; 43(6):240–6. PMID: 12074288.
5. Ogilvie GK, Haschek WM, Withrow SJ, Richardson RC, Harvey HJ, Henderson RA, et al. Classification of primary lung tumors in dogs: 210 cases (1975–1985). J Am Vet Med Assoc. 1989; 195(1):106–8. PMID: 2547742.
6. Moulton JE, von Tscharner C, Schneider R. Classification of lung carcinomas in the dog and cat. Vet Pathol. 1981; 18(4):513–28. PMID: 6266118.
7. Griffey SM, Kraegel SA, Madewell BR. Rapid detection of K-ras gene mutations in canine lung cancer using single-strand conformational polymorphism analysis. Carcinogenesis. 1998; 19(6):959–63. PMID: 9667731.
8. Polton GA, Brearley MJ, Powell SM, Burton CA. Impact of primary tumour stage on survival in dogs with solitary lung tumours. J Small Anim Pract. 2008; 49(2):66–71. doi: 10.1111/j.1748-5827.2007.00403.x PMID: 17725584.
9. Vail DM, Amantea MA, Colbern GT, Martin FJ, Hilger RA, Working PK. Pegylated liposomal doxorubicin: proof of principle using preclinical animal models and pharmacokinetic studies. Semin Oncol. 2004; 31(6 Suppl 13):16–35. PMID: 15717736.
10. Davidson MR, Gazdar AF, Clarke BE. The pivotal role of pathology in the management of lung cancer. J Thorac Dis. 2013; 5 Suppl 5:S463–78. doi: 10.3978/j.issn.2072-1439.2013.08.43 PMID: 24163740; PubMed Central PMCID: PMC3804871.
11. Sakashita S, Sakashita M, Sound Tsao M. Genes and pathology of non-small cell lung carcinoma. Semin Oncol. 2014; 41(1):28–39. doi: 10.1053/j.seminoncol.2013.12.008 PMID: 24565579.
12. Berge EM, Doebele RC. Targeted therapies in non-small cell lung cancer: emerging oncogene targets following the success of epidermal growth factor receptor. Semin Oncol. 2014; 41(1):110–25. doi: 10.1053/j.seminoncol.2013.12.006 PMID: 24565585; PubMed Central PMCID: PMC4159759.
13. Lwin Z, Riess JW, Gandara D. The continuing role of chemotherapy for advanced non-small cell lung cancer in the targeted therapy era. J Thorac Dis. 2013; 5 Suppl 5:S556–64. doi: 10.3978/j.issn.2072-1439.2013.08.47 PMID: 24163748; PubMed Central PMCID: PMC3804876.
14. Alvarez CE. Naturally occurring cancers in dogs: insights for translational genetics and medicine. ILAR J. 2014; 55(1):16–45. doi: 10.1093/ilar/ilu010 PMID: 24936028.
15. Barrott JJ, Haystead TA. Hsp90, an unlikely ally in the war on cancer. FEBS J. 2013; 280(6):1381–96. doi: 10.1111/febs.12147 PMID: 23356585; PubMed Central PMCID: PMC3815692.
16. Taipale M, Jarosz DF, Lindquist S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. Nat Rev Mol Cell Biol. 2010; 11(7):515–28. doi: 10.1038/nrm2918 PMID: 20531426.
17. Saibil H. Chaperone machines for protein folding, unfolding and disaggregation. Nat Rev Mol Cell Biol. 2013; 14(10):630–42. doi: 10.1038/nrm3658 PMID: 24026055; PubMed Central PMCID: PMC4340576.
18. Kamal A, Thao L, Sensintaffar J, Zhang L, Boehm MF, Fritz LC, et al. A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. Nature. 2003; 425(6956):407–10. doi: 10.1038/nature01913 PMID: 14508491.

19. Whitesell L, Lindquist SL. HSP90 and the chaperoning of cancer. Nat Rev Cancer. 2005; 5(10):761–72. doi: 10.1038/nrc1716 PMID: 16175177.

20. Garcia-Carbonero R, Carnero A, Paz-Ares L. Inhibition of HSP90 molecular chaperones: moving into the clinic. Lancet Oncol. 2013; 14(9):e358–69. doi: 10.1016/S1470-2045(13)70169-4 PMID: 23896275.

21. Trepel J, Mollapour M, Giaccone G, Neckers L. Targeting the dynamic HSP90 complex in cancer. Nat Rev Cancer. 2010; 10(8):537–49. doi: 10.1038/nrc2887 PMID: 20651736.

22. Sang J, Acquaviva J, Friedland JC, Smith DL, Sequeira M, Zhang C, et al. Targeted inhibition of the molecular chaperone Hsp90 overcomes ALK inhibitor resistance in non-small cell lung cancer. Cancer Discov. 2013; 3(4):430–43. doi: 10.1158/2159-8290.CD-12-0440 PMID: 23533265; PubMed Central PMCID: PMC4086149.

23. Shimamura T, Perera SA, Foley KP, Sang J, Rodig SJ, Inoue T, et al. Ganetespib (STA-9090), a non-geldanamycin HSP90 inhibitor, has potent antitumor activity in in vitro and in vivo models of non-small cell lung cancer. Clin Cancer Res. 2012; 18(18):4973–85. doi: 10.1158/1078-0432.CCR-11-2967 PMID: 22806877; PubMed Central PMCID: PMC3477583.

24. Socinski MA, Goldman J, El-Hariry I, Koczywas M, Vukovic V, Horn L, et al. A multicenter phase II study of ganetespib monotherapy in patients with genotypically defined advanced non-small cell lung cancer. Clin Cancer Res. 2013; 19(11):3068–77. doi: 10.1158/1078-0432.CCR-12-3381 PMID: 23553849; PubMed Central PMCID: PMC3874465.

25. Supko JG, Hickman RL, Grever MR, Malspeis L. Preclinical pharmacologic evaluation of geldanamycin as an antitumor agent. Cancer Chemother Pharmacol. 1995; 36(4):305–15. doi: 10.1007/BF00689048 PMID: 7628050.

26. London CA, Bear MD, McCleese J, Foley KP, Paalangara R, Inoue T, et al. Phase I evaluation of STA-1474, a prodrug of the novel HSP90 inhibitor ganetespib, in dogs with spontaneous cancer. PLoS One. 2011; 6(11):e27018. doi: 10.1371/journal.pone.0027018 PMID: 22073242; PubMed Central PMCID: PMC3207826.

27. Stokol T, Daddona JL, Mubayed LS, Trimpert J, Kang S. Evaluation of tissue factor expression in canine tumor cells. Am J Vet Res. 2011; 72(8):1097–106. doi: 10.2460/ajvr.72.8.1097 PMID: 21801068.

28. Nemoto Y, Maruo T, Sato T, Deguchi T, Ito T, Sugiyama H, et al. Identification of cancer stem cells derived from a canine lung adenocarcinoma cell line. Vet Pathol. 2011; 48(5):1029–34. doi: 10.1177/0300985810396106 PMID: 21245282.

29. Alvarez CE, Akey JM. Copy number variation in the domestic dog. Mamm Genome. 2012; 23(1–2):144–63. doi: 10.1007/s00335-011-9369-8 PMID: 22138850.

30. Chen WK, Swartz JD, Rush LJ, Alvarez CE. Mapping DNA structural variation in dogs. Genome Res. 2009; 19(3):500–9. doi: 10.1101/gr.083741.108 PMID: 19015322; PubMed Central PMCID: PMC2661804.

31. Mariotti ET, Premanandan C, Lorch G. Canine pulmonary adenocarcinoma tyrosine kinase receptor expression and phosphorylation. BMC veterinary research. 2014; 10:19. doi: 10.1186/1746-6148-10-19 PMID: 24423144; PubMed Central PMCID: PMC3896673.

32. Chari R, Lonergan KM, Pikor LA, Coe BP, Zhu CQ, Chan TH, et al. A sequence-based approach to identify reference genes for gene expression analysis. BMC Med Genomics. 2010; 3:32. doi: 10.1186/1755-8794-3-32 PMID: 20682026; PubMed Central PMCID: PMC2928167.

33. Chou TC, Talalay P. Generalized equations for the analysis of inhibitions of Michaelis-Menten and higher-order kinetic systems with two or more mutually exclusive and nonexclusive inhibitors. Eur J Biochem. 1981; 115(1):207–16. PMID: 7227366.

34. Rathore K, Cekanova M. Animal model of naturally occurring bladder cancer: characterization of four new canine transitional cell carcinoma cell lines. BMC Cancer. 2014; 14:465. doi: 10.1186/1471-2407-14-465 PMID: 24964787; PubMed Central PMCID: PMC4082678.

35. Kent MS, Collins CJ, Ye F. Activation of the AKT and mammalian target of rapamycin pathways and the inhibitory effects of rapamycin on those pathways in canine malignant melanoma cell lines. American journal of veterinary research. 2009; 70(2):263–9. doi: 10.2460/ajvr.70.2.263 PMID: 19231960; PubMed Central PMCID: PMC4169881.

36. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014; 511(7511):543–50. doi: 10.1038/nature13385 PMID: 25079552; PubMed Central PMCID: PMC4231481.

37. Park YK, Kang TW, Baek SJ, Kim KI, Kim SY, Lee D, et al. CaGe: A Web-Based Cancer Gene Annotation System for Cancer Genomics. Genomics Inform. 2012; 10(1):33–9. doi: 10.5808/GI.2012.10.1.33 PMID: 23105926; PubMed Central PMCID: PMC3475486.

38. Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. Cell. 2013; 155(4):948–62. doi: 10.1016/j.cell.2013.10.011 PMID: 24183448; PubMed Central PMCID: PMC3891052.

39. Qi Y, Gregory MA, Li Z, Brousal JP, West K, Hann SR. p19ARF directly and differentially controls the functions of c-Myc independently of p53. Nature. 2004; 431(7009):712–7. doi: 10.1038/nature02958 PMID: 15361884.

40. Bettini G, Marconato L, Morini M, Ferrari F. Thyroid transcription factor-1 immunohistochemistry: diagnostic tool and malignancy marker in canine malignant lung tumours. Vet Comp Oncol. 2009; 7(1):28–37. doi: 10.1111/j.1476-5829.2008.00166.x PMID: 19222828.

41. Yamaguchi T, Hosono Y, Yanagisawa K, Takahashi T. NKX2-1/TTF-1: an enigmatic oncogene that functions as a double-edged sword for cancer cell survival and progression. Cancer Cell. 2013; 23 (6):718–23. doi: 10.1016/j.ccr.2013.04.002 PMID: 23763999.

42. De Luca A, Grassii M, Maiello MR, Di Maio M, Mancino M, De Maio E, et al. Does the sequence of gemcitabine and vinorelbine affect their efficacy in non-small cell lung cancer in vitro? Anticancer Res. 2004; 24(5A):2985–9. PMID: 15517905.

43. Poirier VJ, Burgess KE, Adams WM, Vail DM. Toxicity, dosage, and efficacy of vinorelbine (Navelbine) in dogs with spontaneous neoplasia. Journal of veterinary internal medicine / American College of Veterinary Internal Medicine. 2004; 18(4):536–9. PMID: 15320594.

44. Gaver RC, George AM, Duncan GF, Morris AD, Deeb G, Faulkner HC, et al. The disposition of carboplatin in the beagle dog. Cancer Chemother Pharmacol. 1988; 21(3):197–202. PMID: 3282707.

45. Freise KJ, Martin-Jimenez T. Pharmacokinetics of gemcitabine and its primary metabolite in dogs after intravenous bolus dosing and its in vitro pharmacodynamics. J Vet Pharmacol Ther. 2006; 29(2):137–45. doi: 10.1111/j.1365-2885.2006.00725.x PMID: 16515668.

46. Simon D, Knebel JW, Baumgartner W, Aufderheide M, Meyer-Lindenberg A, Nolte I. In vitro efficacy of chemotherapeutics as determined by 50% inhibitory concentrations in cell cultures of mammary gland tumors obtained from dogs. American journal of veterinary research. 2001; 62(11):1825–30. PMID: 11703031.

47. Knapp DW, Chan TC, Kuczek T, Reagan WJ, Park B. Evaluation of in vitro cytotoxicity of nonsteroidal anti-inflammatory drugs against canine tumor cells. American journal of veterinary research. 1995; 56 (6):801–5. PMID: 7653891.

48. McMahon MB, Bear MD, Kulp SK, Pennell ML, London CA. Biological activity of gemcitabine against canine osteosarcoma cell lines in vitro. American journal of veterinary research. 2010; 71(7):799–808. doi: 10.2460/ajvr.71.7.799 PMID: 20594083.

49. Proia DA, Bates RC. Ganetespib and HSP90: translating preclinical hypotheses into clinical promise. Cancer Res. 2014; 74(5):1294–300. doi: 10.1158/0008-5472.CAN-13-3263 PMID: 24556722.

50. Powers MV, Clarke PA, Workman P. Death by chaperone: HSP90, HSP70 or both? Cell Cycle. 2009; 8 (4):518–26. PMID: 19197160.

51. McCleese JK, Bear MD, Fossey SL, Mihalek RM, Foley KP, Ying W, et al. The novel HSP90 inhibitor STA-1474 exhibits biologic activity against osteosarcoma cell lines. Int J Cancer. 2009; 125(12):2792–801. doi: 10.1002/ijc.24660 PMID: 19544563.

52. Lin TY, Bear M, Du Z, Foley KP, Ying W, Barsoum J, et al. The novel HSP90 inhibitor STA-9090 exhibits activity against Kit-dependent and -independent malignant mast cell tumors. Exp Hematol. 2008; 36 (10):1266–77. doi: 10.1016/j.exphem.2008.05.001 PMID: 18657349; PubMed Central PMCID: PMC3837096.

53. Wu Z, Moghaddas Gholami A, Kuster B. Systematic identification of the HSP90 candidate regulated proteome. Mol Cell Proteomics. 2012; 11(6):M111 016675. doi: 10.1074/mcp.M111.016675 PMID: 22337586; PubMed Central PMCID: PMC3433914.

54. Wen W, Liu W, Shao Y, Chen L. VER-155008, a small molecule inhibitor of HSP70 with potent anti-cancer activity on lung cancer cell lines. Exp Biol Med (Maywood). 2014; 239(5):638–45. doi: 10.1177/1535370214527899 PMID: 24676905.

55. Chatterjee M, Andrulis M, Stuhmer T, Muller E, Hofmann C, Steinbrunn T, et al. The PI3K/Akt signaling pathway regulates the expression of Hsp70, which critically contributes to Hsp90-chaperone function and tumor cell survival in multiple myeloma. Haematologica. 2013; 98(7):1132–41. doi: 10.3324/haematol.2012.066175 PMID: 23065523; PubMed Central PMCID: PMC3696618.

56. Wang Y, McAlpine SR. Regulating the cytoprotective response in cancer cells using simultaneous inhibition of Hsp90 and Hsp70. Org Biomol Chem. 2015; 13(7):2108–16. doi: 10.1039/c4ob02531h WOS:000349401300023. PMID: 25526223

57. Shin HC, Alani AW, Cho H, Bae Y, Kolesar JM, Kwon GS. A 3-in-1 polymeric micelle nanocontainer for poorly water-soluble drugs. Mol Pharm. 2011; 8(4):1257–65. doi: 10.1021/mp2000549 PMID: 21630670.

58. Gorn M, Habermann CR, Anige M, Thom I, Schuch G, Andritzky B, et al. A pilot study of docetaxel and trofosfamide as second-line 'metronomic' chemotherapy in the treatment of metastatic non-small cell lung cancer (NSCLC). Onkologie. 2008; 31(4):185–9. doi: 10.1159/000118626 PMID: 18418020.

59. Budman DR, Calabro A, Kreis W. Synergistic and antagonistic combinations of drugs in human prostate cancer cell lines in vitro. Anticancer Drugs. 2002; 13(10):1011–6. PMID: 12439335.

60. Vinci M, Gowan S, Boxall F, Patterson L, Zimmermann M, Court W, et al. Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. BMC Biol. 2012; 10:29. doi: 10.1186/1741-7007-10-29 PMID: 22439642; PubMed Central PMCID: PMC3349530.

61. Endo H, Okami J, Okuyama H, Kumagai T, Uchida J, Kondo J, et al. Spheroid culture of primary lung cancer cells with neuregulin 1/HER3 pathway activation. J Thorac Oncol. 2013; 8(2):131–9. doi: 10.1097/JTO.0b013e3182779ccf PMID: 23328545.

62. Ekert JE, Johnson K, Strake B, Pardinas J, Jarantow S, Perkinson R, et al. Three-dimensional lung tumor microenvironment modulates therapeutic compound responsiveness in vitro—implication for drug development. PLoS One. 2014; 9(3):e92248. doi: 10.1371/journal.pone.0092248 PMID: 24638075; PubMed Central PMCID: PMC3956916.

63. Kharaishvili G, Simkova D, Bouchalova K, Gachechiladze M, Narsia N, Bouchal J. The role of cancer-associated fibroblasts, solid stress and other microenvironmental factors in tumor progression and therapy resistance. Cancer Cell Int. 2014; 14:41. doi: 10.1186/1475-2867-14-41 PMID: 24883045; PubMed Central PMCID: PMC4038849.

64. Ware KE, Marshall ME, Heasley LR, Marek L, Hinz TK, Hercule P, et al. Rapidly acquired resistance to EGFR tyrosine kinase inhibitors in NSCLC cell lines through de-repression of FGFR2 and FGFR3 expression. PLoS One. 2010; 5(11):e14117. doi: 10.1371/journal.pone.0014117 PMID: 21152424; PubMed Central PMCID: PMC2994708.

65. Lotti F, Jarrar AM, Pai RK, Hitomi M, Lathia J, Mace A, et al. Chemotherapy activates cancer-associated fibroblasts to maintain colorectal cancer-initiating cells by IL-17A. J Exp Med. 2013; 210 (13):2851–72. doi: 10.1084/jem.20131195 PMID: 24323355; PubMed Central PMCID: PMC3865474.

66. Tripathi M, Billet S, Bhowmick NA. Understanding the role of stromal fibroblasts in cancer progression. Cell Adh Migr. 2012; 6(3):231–5. doi: 10.4161/cam.20419 PMID: 22568983; PubMed Central PMCID: PMC3427237.

67. Laederich MB, Degnin CR, Lunstrum GP, Holden P, Horton WA. Fibroblast growth factor receptor 3 (FGFR3) is a strong heat shock protein 90 (Hsp90) client: implications for therapeutic manipulation. J Biol Chem. 2011; 286(22):19597–604. doi: 10.1074/jbc.M110.206151 PMID: 21487019; PubMed Central PMCID: PMC3103339.

68. Jung DW, Kim J, Che ZM, Oh ES, Kim G, Eom SH, et al. A triazine compound S06 inhibits proinvasive crosstalk between carcinoma cells and stromal fibroblasts via binding to heat shock protein 90. Chem Biol. 2011; 18(12):1581–90. doi: 10.1016/j.chembiol.2011.10.001 PMID: 22195560.

Appendix 8

Published article: Canine Mammary Carcinomas: A Comparative Analysis of Altered Gene Expression

Lutful Kabir, Farruk M., Carlos E. Alvarez, and R. Curtis Bird. "Canine Mammary Carcinomas: A Comparative Analysis of Altered Gene Expression." *Veterinary Sciences* 3.1 (2015): 1.

veterinary sciences

MDPI

Review

Canine Mammary Carcinomas: A Comparative Analysis of Altered Gene Expression

Farruk M. Lutful Kabir [1,2], Carlos E. Alvarez [3] and R. Curtis Bird [1,*]

Received: 2 June 2015; Accepted: 21 December 2015; Published: 25 December 2015
Academic Editor: Jaime Modiano

1 Auburn University Research Initiative in Cancer (AURIC), Department of Pathobiology, College of Veterinary Medicine, Auburn University, AL 36849, USA; fzl0005@auburn.edu
2 Current address: Department of Pediatrics, Division of Pulmonology, University of Alabama at Birmingham, Birmingham, AL 35294, USA
3 Center for Molecular and Human Genetics, The Research Institute at Nationwide Children's Hospital Departments of Pediatrics and Veterinary Clinical Sciences, The Ohio State University Colleges of Medicine and Veterinary Medicine, Columbus, OH 43205, USA; Carlos.Alvarez@nationwidechildrens.org
* Correspondence: birdric@auburn.edu; Tel.: +1-334-844-4539; Fax: +1-334-844-2652

Abstract: Breast cancer represents the second most frequent neoplasm in humans and sexually intact female dogs after lung and skin cancers, respectively. Many similar features in human and dog cancers including, spontaneous development, clinical presentation, tumor heterogeneity, disease progression and response to conventional therapies have supported development of this comparative model as an alternative to mice. The highly conserved similarities between canine and human genomes are also key to this comparative analysis, especially when compared to the murine genome. Studies with canine mammary tumor (CMT) models have shown a strong genetic correlation with their human counterparts, particularly in terms of altered expression profiles of cell cycle regulatory genes, tumor suppressor and oncogenes and also a large group of non-coding RNAs or microRNAs (miRNAs). Because CMTs are considered predictive intermediate models for human breast cancer, similarities in genetic alterations and cancer predisposition between humans and dogs have raised further interest. Many cancer-associated genetic defects critical to mammary tumor development and oncogenic determinants of metastasis have been reported and appear to be similar in both species. Comparative analysis of deregulated gene sets or cancer signaling pathways has shown that a significant proportion of orthologous genes are comparably up- or down-regulated in both human and dog breast tumors. Particularly, a group of cell cycle regulators called cyclin-dependent kinase inhibitors (CKIs) acting as potent tumor suppressors are frequently defective in CMTs. Interestingly, comparative analysis of coding sequences has also shown that these genes are highly conserved in mammals in terms of their evolutionary divergence from a common ancestor. Moreover, co-deletion and/or homozygous loss of the INK4A/ARF/INK4B (CDKN2A/B) locus, encoding three members of the CKI tumor suppressor gene families (p16/INK4A, p14ARF and p15/INK4B), in many human and dog cancers including mammary carcinomas, suggested their important conserved genetic order and localization in orthologous chromosomal regions. miRNAs, as powerful post-transcriptional regulators of most of the cancer-associated genes, have not been well evaluated to date in animal cancer models. Comprehensive expression profiles of miRNAs in CMTs have revealed their altered regulation showing a strong correlation with those found in human breast cancers. These genetic correlations between human and dog mammary cancers will greatly advance our understanding of regulatory mechanisms involving many critical cancer-associated genes that promote neoplasia and contribute to the promising development of future therapeutics.

Keywords: canine; mammary cancer; oncogenes; tumor suppressor genes

1. Introduction

In the field of human cancer research, there is an intense interest in development of appropriate model systems for the advancement of future therapeutic inventions. Companion animals such as domesticated dogs (*Canis lupus familiaris*) are considered excellent preclinical models of cancers and other complex human diseases for many reasons, including their easy accessibility and living status in diverse cultures [1]. Since they are treated as pet animals, most of the dog population shares the same environment, risk factors or disease characteristics with the human population [2,3] which provides an added advantage for scientists to investigate cancer etiologies. Additionally, dogs represent a more outbred population than inbred laboratory animals providing a genetic diversity similar to that observed in humans [4].

Canine models address two important issues in cancer research. First, in terms of similarities, dogs spontaneously develop cancers in the context of a natural immune system with a clinical presentation, tumor genetics and heterogeneity, disease progression and response to conventional therapies [5] that better models the complex biology of cancers and their interactions with the immune system in human patients than mouse models. The similarities between the dog and human genomes have also greatly enhanced comparative genomic analysis. With the advent of the high resolution 2.4 billion bp canine genome sequence and the identification of nearly all of its genes as clear orthologs of known human genes [4], the dog has emerged as a valuable comparative and intermediate model for the study of human cancers. The high level of sequence conservation between canine and human genomes are key to this comparative analysis especially since 600 Mb of DNA sequence conserved between dog and human is missing from the murine genome [1]. Secondly, using dogs as animal models may contribute to the development of cancer therapeutics for, not only human and dog, but also other species—a promising theme lately coined as "One Medicine" that campaigns under a unified scientific platform where discoveries in one species can be translated to others to improve health management in all species. Canine tumors with potential relevance for human cancer biology include osteosarcoma, mammary carcinoma, lymphoma, melanoma, lung carcinoma, and soft tissue sarcomas [6].

2. Canine Mammary Tumors (CMTs)

Mammary tumors are the most common neoplasm in sexually intact female dogs. The severity of canine mammary tumors (CMT) can be appreciated from a number of studies that reported increased rates of incidence in the dog population globally. Breast cancer represents the second most frequent neoplasm in humans and all dogs after lung and skin cancers, respectively, although many reports indicate that dogs are two to four times more susceptible to mammary cancers than women in certain geographical areas [7–10]. Nearly 50% of all these neoplasms are diagnosed as malignant and more than 95% of these malignant CMTs are carcinomas [11,12].

Canine mammary carcinomas are biologically heterogeneous neoplasms offering several ways to classify such tumors on the basis of histopathological characteristics or expression of molecular markers [13]. Despite the appearance of histomorphological variations between human and canine breast cancers, due to various prognostic indicators, a number of studies have reported that there are significant similarities regarding molecular marker expression, hormone dependency and cancer phenotypes [11–14]. It is important to classify breast cancer in order to correlate clinical phenotypes, invasion or grade of progression and to develop prognostic markers. The human classification of breast cancers based on expression profile of luminal epithelial specific genes and hormone receptors including estrogen receptor 1 (ESR1), progesterone receptor (PR) and proto-oncogenes such as epidermal growth factor receptors (EGFR/HER2), have also identified similar molecular subtypes in CMTs, but unlike human subtypes, these are not routinely investigated for CMTs during clinical diagnosis [15,16]. Recently, in more refined studies employing immunohistochemical approaches and based on the characteristic expression patterns of ESR1, PR and EGFR (ERBB1/HER1, ERBB2/HER2, ERBB3 and ERBB4), human-like breast cancer phenotypes for CMTs have been developed

and classified as luminal A, luminal B, HER2 positive and triple negative (basal-like) [17,18]. Such standard classification therefore strongly supports canine mammary tumors as valuable intermediate models for human breast cancer that should be well-placed for developing diagnostic and treatment strategies.

Because CMTs are considered predictive models for human breast cancer [6], similarities in genetic alterations and cancer predisposition between humans and dogs have raised interest even further. A large number of studies have demonstrated that CMTs have many similarities in molecular and clinical features with human breast cancer. Many genetic/epigenetic/tumor biology traits that are most frequently associated with mammary cancer have been identified and comparative gene expression analysis has revealed a significant similarity in the canine and human genes associated with mammary tumor development [19]. Although CMTs have not yet been classified based on surface markers, due to an absence of appropriate antibodies identifying human breast cancer subtypes, the expression profile of vital genes involved in cellular proliferation, angiogenesis, apoptosis, cell cycle regulation, DNA damage repair, signal transduction, and survival pathways firmly correlate to those in human breast cancer [19,20]. These studies characterized CMTs, based on genome-wide gene expression changes, comparing to human breast cancer, suggesting that mutations and alterations in the cancer genome may promote deregulation of individual genes in mammary cancers.

Comparative analysis of deregulated gene sets or cancer signaling pathways showed that a significant proportion of orthologous genes are comparably up- and down-regulated in both human and canine breast tumors. Prominent oncogenic pathways and related genes, such as PI3K/AKT, KRAS, MAPK, Wnt, β-catenin, BRCA2, ESR1 and P-cadherin, are commonly up-regulated while representative tumor suppressive pathways, such as p53, p16/INK4A, PTEN and E-cadherin, are down-regulated in human and canine breast cancer [19,21–25]. This chapter will discuss the comparative aspects of cell cycle regulatory genes, particularly the evolutionary descent, structure, genomic localization, biological functions, expression defects and post-transcriptional regulation of the CDKN2/INK4 family of cyclin-dependent kinase inhibitors (CKIs) in canine and human breast cancers.

3. Cell Cycle Regulators: A Classic Repertoire of Tumor Suppressors

From simple eukaryotes such as yeast to higher mammals, the cell cycle serves as a fundamental biological process by which cells grow and divide and its regulation is central to cancer promoting mechanisms. Cancer causes are complex involving dysregulation of cellular functions and dysregulation in micro-environmental signaling as well as being rooted in oncogenic mutations [26]. Ultimately, cancers occur due to an alteration in the regulation of cell proliferation. Cell proliferation itself is rooted in the cell cycle which is a highly regulated process governed by complex mechanisms [27]. The uncontrolled cell proliferation in cancer is also associated with a vicious cycle where cells divide through unchecked cell cycle progression with a reduction in sensitivity to signals that normally guide cells to adhere, become quiescent, terminally differentiate or die. This combination of unregulated proliferation and a failure of balancing suppressor activities is hallmark of malignant transformation resulting in neoplasia that can eventually develop the ability to spread and migrate throughout the body through metastasis [28]. One such group of genetic alterations that contribute to cancer development are often termed hypermorphic mutations that largely define oncogenes and result from the mutated versions of normal cellular proto-oncogenes. Oncogenic mutations appear to destroy the integrity and modulated control of cell proliferation first by altering control of the stimulatory pathways that promote cell growth. They may also promote neoplasia by suppressing those pathways normally responsible for modulating and inhibiting proliferation or causing exit from the cell cycle entirely. These “loss of function” mutations occur in tumor suppressor genes that encode proteins that can negatively regulate cell cycle progression but, when mutated, are permissive for cancer development and can promote spontaneous as well as,

in some instances, hereditary forms of cancer [29]. Two important examples of such loss-of-function mutations affecting cell cycle regulation are mutations in the retinoblastoma (Rb) and p16/INK4A tumor suppressor genes [30,31]. Loss of function of these tumor suppressor gene products results in liberation of the E2F transcription factors, associated with S phase promotion, that consequently remove control of cell cycle exit during G1 phase resulting in abnormal and continuous cellular proliferation [29].

A group of inhibitory proteins, called cyclin-dependent kinases inhibitors (CKIs) or CDK inhibitors, control cyclin-CDK activity thereby restraining cell cycle progression in response to extracellular and intracellular signals [32,33]. The orderly progression of the cell cycle is fine-tuned by the genes encoding such negative regulators, or CKIs, and positive regulators including the cyclins and CDKs. Dysregulation of these genes can lead to premature and unregulated entry into the next phase of the cell cycle leaving the previous phase unchecked, and frequently this occurs prior to completion of critical molecular events such as repair of DNA damage or replication errors. Such dysfunction frequently triggers genomic instability and neoplastic transformation. Based on their structural similarities and specific roles in cell cycle regulation, CKIs are divided into two distinct groups: the INK4, or CDKN2, and the Cip/Kip, or CDKN1, families [33]. The first group representing the INK4 proteins (Inhibitors of CDK4) are so named because of their ability to specifically inhibit the catalytic subunits of CDK4 and CDK6. It has been reported that INK4 proteins compete with cyclin D for binding to the CDK4/6 subunit [34,35]. The members of the INK4 protein family that share common structural features are p16/INK4A (and p14ARF, an alternatively spliced product from the same locus), p15/INK4B, p18/INK4C and p19/INK4D (Figure 1) [36–38]. The Cip/Kip family (for CDK interacting protein/ Kinase inhibitory protein) consists of three members, including p21/Cip1, p27/Kip1 and p57/Kip2, all of which share a common inhibitory domain that enables them to bind CDK complexes [37,39]. These proteins of the Cip/Kip family have broad specificity for binding and inhibiting a number of cyclin-CDK complexes compared to that of INK4 members. They also inhibit the activity cyclin D-CDK4 preventing Rb phosphorylation during G1 to S phase transition. In addition, they inhibit cyclin A-CDK2 in late G1 phase and cyclin E-CDK2 in early S phase (Figure 1) [37]. Therefore, both CKI families are important modulating components of the complex network of cell cycle regulatory mechanisms.

3.1. CDK Inhibitors Form a Repertoire of Tumor Suppressor Proteins

Many studies stress the fact that CDKs are positive regulators and CKIs are negative regulators of cell proliferation based on their distinct inhibitory actions in the eukaryotic cell cycle [38,40]. Besides their specific roles in cell cycle regulation, differentiation and development, CKIs are proven or highly likely tumor suppressors to have this potential, as mutations in these genes promote malignant phenotypes [32,36,37,41]. In some clinical trials, CKI tumor suppressors aggressively promote cancer cell growth by inducing p53 function and stability and increasing anti-proliferative activity thereby inhibiting cell cycle progression [42]. Among all the CKIs, p16/INK4A is the founding member and was the first classified as a major tumor suppressor gene (only preceded by p53 for many human malignancies) because the mutations in the INK4A/ARF locus, and loss of heterozygosity of the chromosomal region encoding this gene, have been reported in a wide range of cancers including melanomas, leukemias, gliomas, lung, breast and bladder cancers [36,41]. The loss of expression of the neighboring p15/INK4B gene, due to promoter hypermethylation, also occurs in a number of leukemias and lymphomas [38,43]. The p16/INK4A locus has also been found to be frequently mutated in canine malignant melanomas, mammary carcinomas and fibrosarcomas [21,44–47].

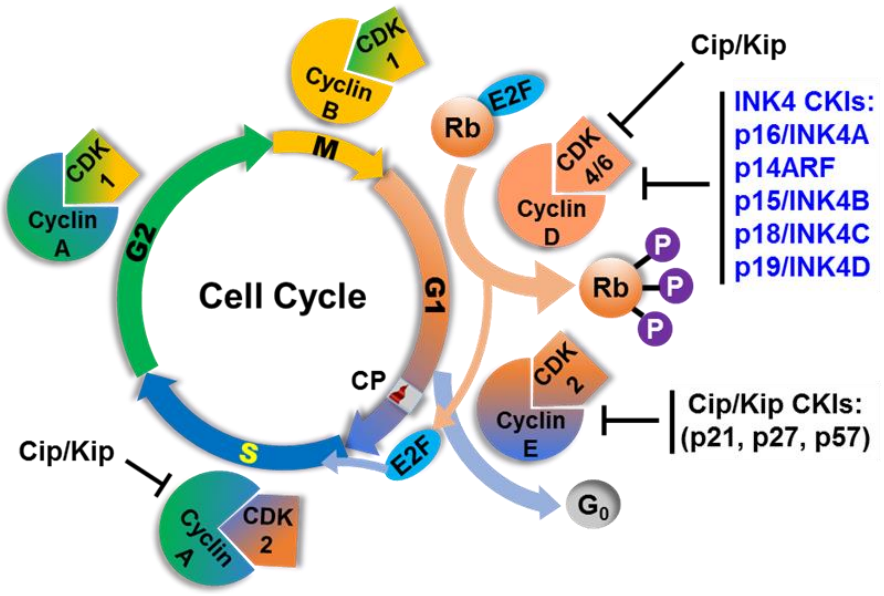


Figure 1. Cyclin-Dependent Kinase Inhibitors (CKIs) are regulators of the cell cycle. Cell cycle phases, major regulatory proteins or protein complexes including cyclins, CDKs, INK4 and Cip/Kip inhibitors and their targets are shown. Checkpoint (CP) or major restriction point.

3.2. Evolutionary History, Genomic Localization and Structure of the INK4A/ARF Locus

A locus on the short arm of human chromosome 9p21 (a known multiple tumor suppressor locus) encodes three products called p16/INK4A, p14ARF and p15/INK4B, all of which regulate cell proliferation by inhibiting the cyclin-CDK complex at the G1 to S phase transition in the cell cycle [36,41]. The organization of the INK4A/ARF/INK4B (or INK4A/ARF) locus in the mammalian genome is highly conserved. Orthologous sequence searches and comparative genomics analysis has demonstrated that this locus in human (chromosome 9) is syntenic to that of chimp (chromosome 9), dog (chromosome 11), cat (chromosome D4), mouse (chromosome 4) and rat (chromosome 5) and this region, encoding several tumor suppressor genes, is highly susceptible to genetic instability and mutations in many cancers [36,37,48]. The close similarities between the INK4A and INK4B genes and two other members of the INK4 CKI gene family, INK4C and INK4D based on their protein sequence, biochemical properties and functions in the cell cycle, suggest that they arose as a result of gene duplication during the course of evolution. This is most likely true since a number of studies have demonstrated that all four INK4 CKIs share a common structural feature called ankyrin repeats that appear to function as a structural scaffold facilitating protein–protein interactions and these four CKIs also appear functionally related [36,49,50].

The evolutionary history of the INK4A/ARF/INK4B locus suggests that the INK4 genes have evolved through tandem gene duplication events. One of the most interesting findings for the evolutionary descent of INK4 genes was the complete absence of ARF-like gene products in the Japanese puffer fish *Fugu rubripes* (fugu) and in the zebrafish [37,48] suggesting that p14ARF was introduced into the vertebrate or mammalian genome following INK4 duplication. Three unique INK4 genes, representing INK4A or B, INK4C and INKD have been identified in the fugu genome (Figure 2). Evolutionarily, p16/INK4A and p15/INK4B are products of a local tandem duplication while p18/INK4C and p19INK4D are present on other chromosomes [37,48]. Cross-species comparative analysis suggested that a single common ancestral INK4 gene was present and a series of duplication and rearrangement events first gave rise to INK4A/B and INK4C/D-like elements in a common vertebrate ancestor and after the divergence of higher vertebrates from tetrapod and fish approximately 350 million years ago (MYA) gave rise to the individual INK4 genes in the mammalian genome [37].

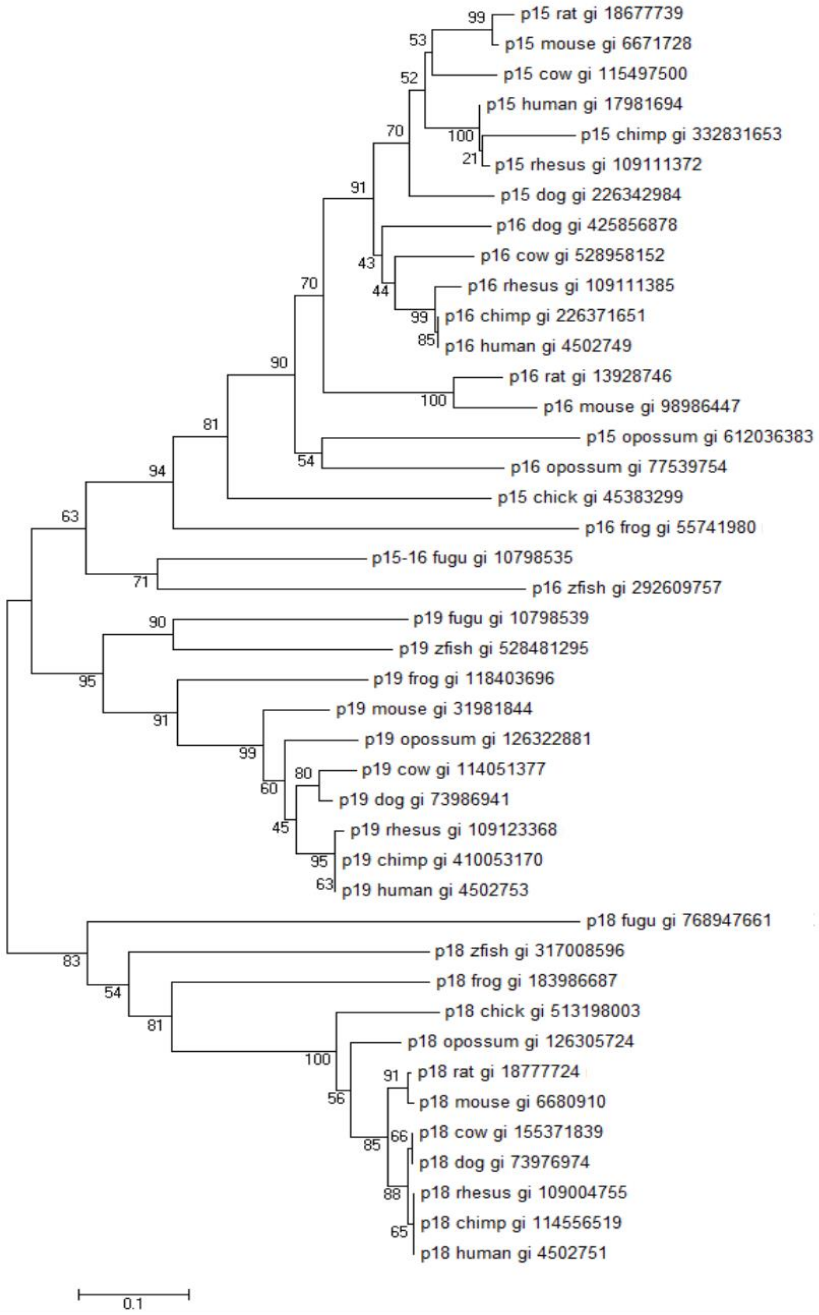


Figure 2. INK4/CDKN2 family tree. Annotated INK4 proteins from select organisms were aligned using Clustal W. The alignment was used to construct a neighbor-joining phylogenetic tree (applying complete deletion of gaps and Poisson model rates and patterns; MEGA6). Bootstrap values were calculated from 500 repetitions. Similar results were achieved with maximum parsimony phylogenetic treeing (not shown). The phylogenetic analyses (see text) demonstrate the high similarity and conservation among INK4 proteins as well as their evolutionary descent. The scale bar shows the number of substitutions per site. NCBI GI accession numbers of the proteins are given on the tree along with the common or abbreviated animal name. Taxonomy abbreviations follow: zfish, *Danio rerio* (zebrafish); fugu, *Takifugu rubripes* (Japanese puffer fish); *Xenopus tropicalis* (western clawed frog); chick, *Gallus gallus* (chicken, red junglefowl); opossum, *Monodelphis domestica*; cow, *Bos taurus*; dog, *Canis lupus familiaris*; rat, *Rattus norvegicus* (Norway rat); mouse, *Mus musculus* (house mouse); rhesus, *Macaca mulatta* (Rhesus macaque); chimp, *Pan troglodytes* (common chimpanzee); human, *Homo sapiens*.

These evolutionary changes placed p16/INK4A and p15/INK4B about 30 kb apart in the same transcriptional orientation on chromosome 9p21 whereas p18/INK4C and p19/INK4D are present on human chromosomes 1p32 and 19p13, respectively [36]. Phylogenetic trees based on the published amino acid sequences of INK4 proteins indicates their high similarities among groups and likely divergence from a common ancestor (Figure 2). This evolutionary relationship suggests that the p16/INK4A and p15/INK4B from mammals represent a paralogous group that was once related to p16/15 in fugu or zebrafish while p18/INKC and p19/INK4D are more closely related to corresponding orthologs in fugu. Figure 2 suggests that the complement of INK4 genes arose before the marsupial-placental mammal divergence.

During the evolution of p16/INK4A and p15/INK4B through gene duplication, an additional exon appeared when comparing the two genes. This alternative exon is designated exon 1β which is alternatively spliced to exon 2 and 3 of p16INK4A making the novel p14ARF transcript (Figure 3) [37]. Previously, it was postulated that exon 1β was the original exon 1 of the INK4A locus but later it was determined that this alternative exon was transcribed from its own separate promoter and not from the promoter of p16/INK4A exon 1α [51]. The presence of such a separate promoter for p14ARF suggests that its transcription is regulated independently of p16/INK4A. Gene duplication, rearrangement and deletion appear to have resulted in a duplicated exon 1β located in the intergenic region between the INK4A and B genes that later diverged from each other [37,48]. Interestingly, the highly cancer resistant naked mole rat has recently been shown to have an unusual fusion of the p15/p16 tumor suppressor genes (PMID: 25550505). This may represent a further gene rearrangement whose relationship to relative cancer resistance and unusually long life in this rodent remain to be investigated.

3.3. Roles of INK4A/ARF Encoded Regulators in the Cell Cycle and Cancer

The existence of p16/INK4A protein was first discovered as a binding partner of cyclin D-dependent CDK4 by the co-immunoprecipitation assay. In cells transformed by SV40 virus, CDK4 was found to be predominantly associated with p16 rather than cyclin D unraveling an important function of this founder member of the INK4 family and suggesting that p16 can directly bind to the catalytic CDK4 subunit in the absence of regulatory cyclin D [40]. Other INK4 members (p15, p18 and p19) were found to interact with CDK4 and CDK6 by two-hybrid screening. Both *in vitro* and *in vivo* studies have reported that all of the four INK4 proteins directly bind the kinase subunits (CDK4/6) rather than the cyclin subunit (cyclin D) as they act as competitive inhibitors of the cyclins [52]. This specific interaction with CDKs distinguishes the INK4 family from the Cip/Kip family of CKIs [36]. Because there is no sequence similarity between exon 1β of p14ARF and exon 1α of p16 and alternative splicing of exon 1β to the shared exon 2 allows translation to continue from the −1 nucleotide of the open reading frame of p16, p14ARF encodes a completely different protein compared to p16. These two proteins also function in distinct biological pathways. Rb is a critical substrate for cyclin D-dependent kinases [40,53] and its phosphorylation is required to release and activate the E2F transcription factors switching on gene expression involved in the G1 to S phase transition [54]. p16/INK4A and the three other INK4 members prevent Rb phosphorylation by inhibiting CDK4/6 binding with cyclin D [34,35]. This cascade pathway in turns leads to E2F repression that inhibits the transcription of many genes required for exit from G1 and initiation of S phase eventually resulting in growth arrest [37,54].

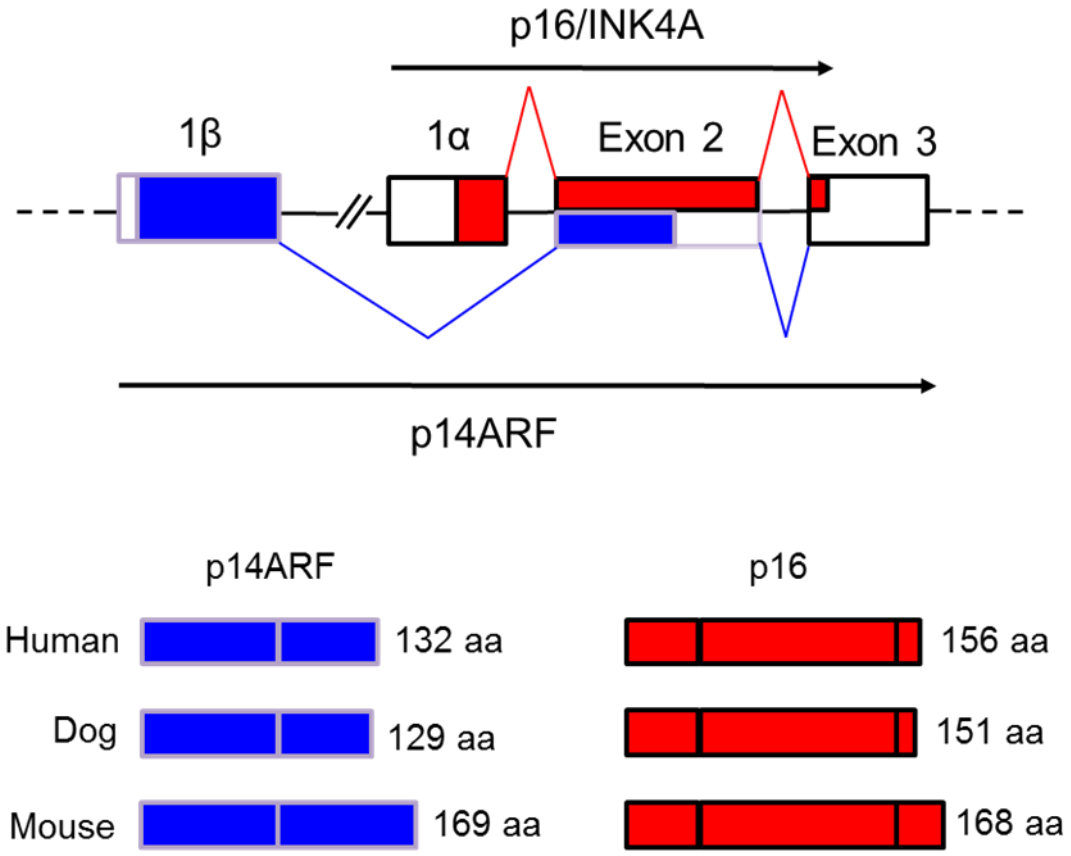
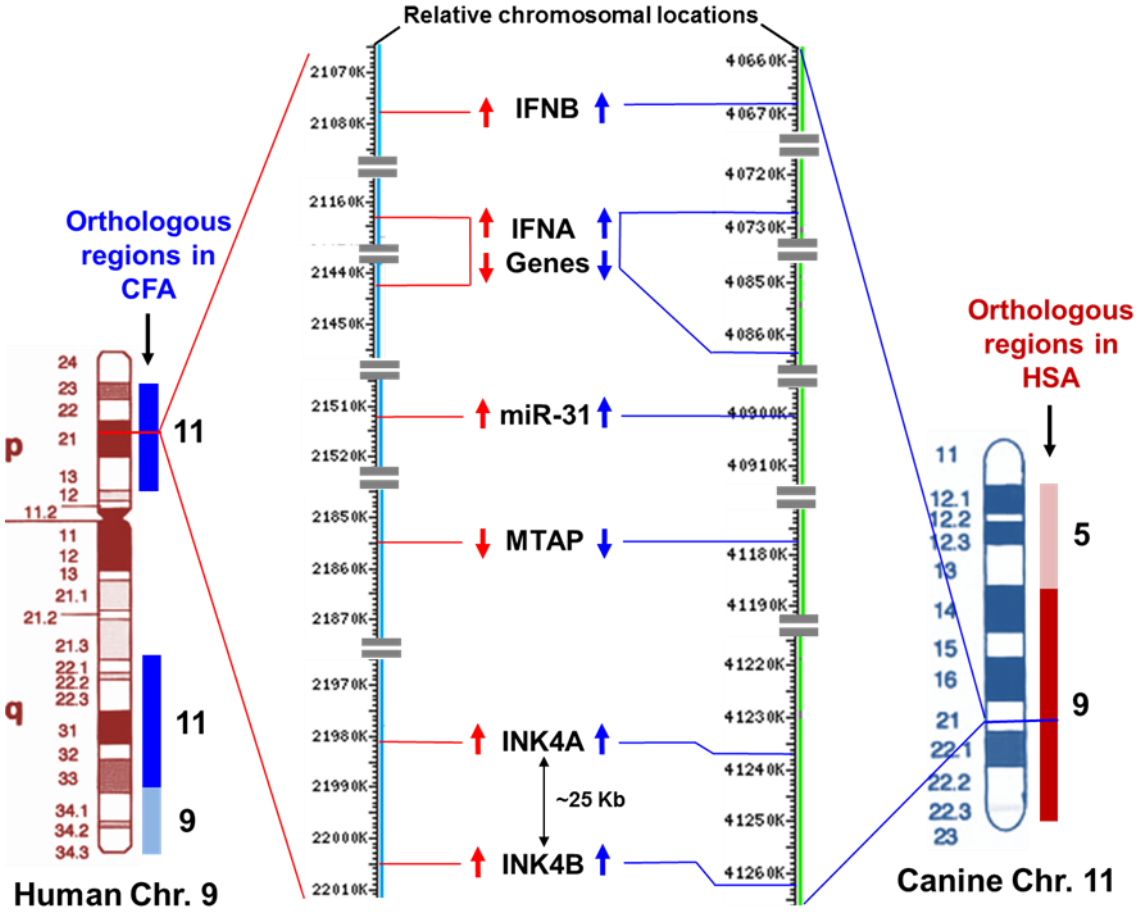


Figure 3. Alternative splicing results in two different transcripts and protein products from the modern INK4A/ARF locus. The exons are shown as boxes and the sequences encoding p16/INK4A are shown as red shading while those encoding the ARF transcript are colored blue. Exon 1α is spliced to INK4A exon 2 and 3 forming the p16 mature transcript whereas exon 1β is alternatively spliced to the same exon 2 and 3 generating the mature p14ARF transcript. The latter produces a different protein from p16 because translation occurs from an alternative reading frame. The sizes of the respective human, dog and mouse p16 and ARF proteins are shown in the bottom panel. (p14ARF in mouse is named p19ARF due to its increased length but should not be compared to p19/INK4D).

On the other hand, p14ARF is highly unlikely to act as a direct inhibitor of CDK4/6 because of its structural differences from other INK4 proteins. A great number of studies, using mouse models and human cancer cells, differentiated the functions and regulation of p14ARF from that of p16. The initial evidence for its anti-proliferative role came from observations that expression of p19ARF (the p14ARF ortholog in mouse) in embryonic fibroblasts or NIH 3T3 cells induced cell cycle arrest but no direct interaction with CDK complexes was detected in immunoprecipitation assays [55]. It has been reported that loss of p19ARF obviates the requirement of p53 inactivation to immortalize mouse embryonic fibroblasts and tumors, including melanomas, *in vivo* [56,57]. This understanding was further refined by other studies demonstrating that suppression of oncogenic transformation in primary cells by p19ARF is abrogated when p53 is inactivated by viral oncoproteins or dominant p53 mutants [58] implying that p19ARF functions upstream of the p53 pathway. Moreover, some groups reported that p19ARF can associate with MDM2 (a p53 ubiquitin protein ligase) or inhibit the E3-ligase activity of MDM2 to prevent MDM2-induced p53 degradation [57–61] suggesting that these proteins—p19ARF, MDM2 and p53—exist in a common regulatory pathway. In addition to p53 stabilization, p14ARF regulates p53 transactivation activity. p53 normally acts as a strong transcriptional activator of p21/Cip1 protein [38]. Expression of p19ARF in primary mouse cells expressing functional p53 results in the induction of p21 that plays essential roles in

G1 to S phase arrest, apoptosis and tumor growth suppression [57–59]. Investigating mutations and gene expression profiles of cell cycle regulatory proteins in many human cancer cell lines and primary tumors provided evidence that p53 mutations do not directly correlate with either p16 or Rb expression [30] stressing the fact that p14ARF (in the p53 pathway) and p16 (in the Rb pathway) have distinct or non-overlapping, important biological functions in cell cycle regulation and cancers [36,37]. Thus, this single locus capably regulates the key pathways controlling cell proliferation—the Rb-dependent and p53-dependent pathways.



Frequently deleted regions in human chr. 9p21 and orthologous canine chr. 11

Figure 4. Relative molecular and cytogenetic mapping of the INKA/ARF locus and closely related genes with their positions on human and canine chromosome 9 and 11, respectively. The regions at human chromosome 9 and canine chromosome 11 that are frequently deleted in cancers are completely orthologous to each other. The molecular mapping shows the exact chromosomal position of these genes extrapolated from the NCBI map view of each chromosome represented by the current human and canine annotation from releases 106 and 103, respectively. The red and blue arrows indicate the transcriptional orientation of genes in the human and dog chromosomes, respectively. Transcription of genes from the "+ strand" is indicated by down arrows and from the "− strand" by up arrows. (CFA = Canis lupus familiaris; HSA = Homo sapiens; Chr. = Chromosome).

3.4. Alteration of the INK4A/ARF Locus in Human and Canine Cancers

There is compelling genetic evidence from numerous cancer studies that p16/INK4A is a critical tumor suppressor gene whose direct inactivation by point mutation, deletion, or promoter hypermethylation is observed in nearly one third of human cancers, establishing its loss as one of the

most frequent lesions promoting human malignancy [37]. The p16/INK4A gene was independently isolated as a candidate tumor suppressor gene located at human chromosome 9p21, the region which is highly conserved across mammals, and was found to be frequently deleted in many human tumors and linked to hereditary susceptibility to melanoma [60–62]. The emergence of human chromosome 9p21 as a site of a major tumor suppressor gene was deduced from extensive cytogenetic and loss of heterozygosity (LOH) studies on a wide range of tumors such as leukemias, melanomas, gliomas, pancreatic adenocarcinomas, as well as breast, lung and bladder cancers [61,63–68]. LOH of chromosome 9p21 that encodes the INK4/ARF locus was also deleted in the study of a neighboring gene called methylthioadenosine phosphorylase (MTAP) that also mapped to the same chromosomal region [69]. MTAP, a regulatory gene for purine and polyamine biosynthesis, is frequently deleted in different malignant cancer cell lines that also have homozygous deletion of p16 suggesting that loss of MTAP in malignant cells is primarily due to linkage between the MTAP and p16 genes on the same chromosomal region and so they were co-deleted [70]. Furthermore, some malignant cells were found to have homozygous deletion of p16 and MTAP but retained an intact p15 gene. These findings of homozygous deletion of p16 and its neighbor in cancer cells also revealed the gene order on chromosome 9p21 starting from the centromeric end which is p15, (p14ARF) p16, MTAP, IFNA and IFNB (interferon alpha and beta) (Figure 4) [70].

Studies with non-human animal models of cancers have also reported genetic defects in the INK4A/ARF locus. The *in vivo* role of p16 in tumorigenesis was initially indicated from mapping tumor susceptibility alleles in common BALB/c mouse strains. This mouse model is prone to tumor development such as plasmacytoma (tumors of the plasma cells) and lung adenocarcinoma in which the major genetic determinant responsible for a strong cancer predisposition also mapped to the INK4A/ARF locus [71,72]. Mice with targeted deletions of p16, p19ARF or both were investigated by several groups suggesting that mouse strains with specific inactivation of p16 or p19ARF were tumor prone but neither genetic loss alone was as severe as those with double knockouts of both of these genes [57,73–75]. Mutant mice that were deficient for p16 and heterozygous for p19ARF spontaneously develop a wide range of tumors including melanoma [95]. Importantly, primary melanomas, mammary carcinomas and osteosarcomas from dogs have also been reported to harbor frequent defects in p16/INK4A [21,45,46,76]. Altered expression profiles from p16/INK4A/ARF have been recurrently observed in a number of canine breast cancers, melanomas and other primary tumors that highly correlate to lesions in humans and mice [30,37,77]. In fact, altogether, deletions or point mutations causing shifting of reading frame and altered expression located mostly in exon 1α have been found in cancers from humans, dogs and mice suggesting that specific mutation mapping in p16/INK4A and its regulation are not limited to these cancer types (for example, melanomas and breast cancers) but also occur in other tumors commonly encountered in mammalian species with neoplasms or uncontrolled cellular growth [21,71,72,77,78].

Furthermore, the region at canine chromosome 11 (orthologous to human chromosome 9p21) encoding INK4A/ARF, MTAP and close neighbors including miR-31, as shown in the comparative chromosomal mapping (Figure 4), is also highly susceptible and prone to concomitant deletion in many cancers in dogs [16]. Reports from several studies suggested that a haplotype spanning MTAP and INK4A/ARF loci showed susceptibility to naturally occurring canine sarcomas [79]. The miR-31, one of the highly cited tumor suppressor miRNAs in human breast cancer, was also found to be down-regulated and differentially expressed in canine osteosarcoma and mammary tumors, respectively [80]. Therefore, the comparative analysis of cytogenetic and molecular mapping of the genetic defects at human chromosome 9p21 and the corresponding canine chromosome 11 identified frequently deleted regions encoded by the INK4A/ARF/INK4B locus with a highly conserved order of genes (Figure 4) that are concurrently lost in many cancers recapitulating the strong similarities in genetic alternations and cancer predisposition between humans and dogs.

4. Regulatory, Small Non-Coding RNAs: microRNAs in Cancers

It is increasingly apparent that a significant portion of the mammalian genome (estimated to be >70%) encodes regulatory information that is largely carried out by non-coding RNAs [81–83]. These non-coding RNAs consist of two major classes: small non-coding RNAs (<200 bp, including miRNAs) and long noncoding RNAs or lncRNAs (>200 to ~100 kb) [84,85]. lncRNAs share many features of mRNAs, but in contrast to mRNAs, they are found within introns of protein coding genes or intergenic regions of the genome [83] demonstrating developmental and tissue-specific expression patterns [86]. The lncRNAs play a number of important regulatory functions that affect epigenetic changes including chromatin remodeling, transcriptional co-activation and repression, post-transcriptional modification of mRNAs as well as cellular functions including differentiation and homeostasis [84]. Dysregulated expression of lncRNAs causes disruption of these biological functions and plays a critical role in cancer development [85]. To date, a number of lncRNAs have been implicated in breast cancer development and metastasis. One of the most well-known and first identified lncRNAs is a HOX antisense intergenic RNA that is commonly abbreviated as *HOTAIR*. This lncRNA, located in the mammalian HOXC locus, has been demonstrated to be associated with polycomb repressive complex 2 that mediates transcriptional repression of numerous genes involved in differentiation pathways during development and stem cell pluripotency [87–90]. Importantly, *HOTAIR* has been reported to be highly upregulated in both primary and metastatic breast cancers and its overexpression is a strong predictor of metastasis and poor survival [87]. However, unlike the rapid advances in miRNA research, including the well-established mechanisms of miRNAs in gene silencing and the strong sequence conservation of miRNAs across mammals, knowledge regarding the molecular mechanisms of lncRNA function in cancer is still growing. Most lncRNAs are poorly conserved, and their mechanisms of action remain unclear and are in need of further exploration [83,84].

The discovery of microRNAs (miRNAs) established a new era in translational regulation research and for understanding post-transcriptional regulation of genes as well as their critical regulatory roles in diverse biological processes including cell cycle, cell proliferation, differentiation, development and apoptosis as well as in disease pathogenesis [91–93]. miRNAs are evolutionarily conserved, endogenous small structural RNA molecules (~22 nucleotides) that post-transcriptionally suppress gene expression in a sequence specific manner [94]. Expression of these small structural RNAs is tightly regulated during development and in normal mature tissues and is frequently altered in cancer [95]. Strikingly, more than 50% of miRNA genes are located in cancer associated genomic regions or fragile sites that are also preferential sites for translocation, deletion, amplification, and integration of exogenous genome fragments suggesting that miRNAs play an important role in the pathogenesis of many human cancers [96,97]. Since miRNAs are encoded by highly conserved and naturally occurring genes across mammalian species, evaluation of their expression profiles in cancer models shows great promise for advancing the development of future therapeutic reagents, as well as for improving diagnostic and prognostic analysis.

miRNAs and their associated proteins appear to be one of the most abundant biomolecules in the cell. Improvements in small nucleotide amplification technologies and in sequence prediction algorithms, miRNA discovery from model organisms, as well as non-model species, have greatly advanced with 35,828 mature miRNAs in 223 species putatively identified to date (miRBase v20.0). Based on this latest estimate, the human, mouse and canine genomes account for 2588, 1915 and 453 mature miRNAs, respectively, and these numbers reflect those miRNAs identified to date and this is anticipated to grow to similar numbers in most mammals compared to the human genome. The number of experimentally validated miRNAs from each species is smaller than the predicted number. However, both bioinformatics and empirical evidence suggests that more than 30% of protein-coding genes in the human genome are subjected to regulation by miRNAs, indicating their prominence as global regulators of gene expression [98–101].

In mammals, mature miRNAs generated from sequential processing of primary miRNA transcripts by Drosha and Dicer miRNA processing complexes associate with 3′untranslated regions (3′UTR) of specific target messenger RNAs (mRNAs) to suppress translation and may also induce their degradation [102]. In the nucleus, the RNase III-type enzyme Drosha processes the long primary transcripts (pri-miRNA that is initially transcribed by RNA polymerase II from the cellular genome), yielding 60–70 nucleotide hairpin precursors called pre-miRNA. The resulting pre-miRNA hairpins are translocated to the cytoplasm by Exportin-5. In the cytoplasm, the pre-miRNAs are further cleaved and processed into 19–25 nucleotide miRNA duplex structures by the RNase Dicer and transactivator RNA binding protein (TRBP). The functional strand (or guide strand) of the mature miRNA is loaded together with Argonaut (Ago2) proteins into an RNA-induced silencing complex (RISC), where it guides RISC to silence target mRNAs through mRNA cleavage, translational repression or deadenylation. The passenger strand (the complementary strand of the double stranded pre-miRNA following Dicer processing) is typically degraded [102]. The mature miRNAs usually target the 3′UTR of mRNAs and make complementary base pairing with their seed (core orthologous target) sequences (located at 2–8 bases from the 5′end of the miRNA) [97]. The seed sequence, by which miRNAs bind to their targets, is only several nucleotides long, suggesting that each miRNA may potentially bind to a large number of genes thereby regulating their expression. miRNAs can direct the RISC complex to downregulate target gene expression by either of two post-transcriptional mechanisms: mRNA cleavage or translational repression [98,103,104]. The execution of one of these mechanisms is primarily determined by the degree of complementarity between the miRNA and its target mRNA. The miRNA will promote the cleavage of the target message if its seed region is sufficiently complementary to the target sequences [105]. After degradation of the mRNA, the miRNA remains intact and can guide the RISC to target other messages. Interestingly, miRNAs can regulate their own expression or biosynthesis by targeting the miRNA processing machinery. For example, the miR-103/107 family can inhibit DICER expression and induce epithelial to mesenchymal transition (EMT) promoting metastasis in human breast cancer [106].

4.1. OncomiRs: Cancer Associated miRNAs

The association of miRNAs with the initiation, progression and key control pathways of human malignancies holds great potential for new developments in advanced diagnostic and therapeutic strategies in the management of most common cancers. The expression of miRNAs are deregulated in cancer by a variety of mechanisms including amplification, deletion, mutation or epigenetic silencing [107–109]. Epigenetic regulation of miRNAs is mediated by promoter hypermethylation in certain human cancers. For example, miR-127, which is downregulated in human cancer cells, has been reported to be located within a CpG island and highly up-regulated by DNA demethylation and histone acetylation [109]. Many groups have discovered "miRNA signatures" in both hematological and solid tumors that discriminate cancers from normal cells and have potential for improving prognosis, management of progression and possibly suppression of cancer [97,110–114]. miRNAs are often regarded as "oncomiRs" meaning miRNAs involved in dominant cancer regulatory mechanisms. OncomiRs can be categorized as tumor oncogenes and tumor suppressors as anti-oncomiRs. miR-155 was one of the first identified oncomiRs that has been demonstrated to be highly expressed in several well-known lymphomas, leukemias, breast, colon and lung cancers [111,113–116]. Like miR-155, other oncogenic miRNAs usually target tumor suppressor genes and cell cycle inhibitors, or other anti-proliferative genes and they can also serve as potential therapeutic targets. Another strong oncogenic candidate miRNA is miR-21 which is upregulated in a wide variety of blood related and solid tumors including myeloid leukemia, lymphocytic leukemia, gliobalstoma and cancers of the pancreas, prostate, stomach, colon, lung, liver and breast [110,113,117–119]. Overexpression of miR-21 in these cancers inhibits the apoptotic pathway promoting dysregulated cell proliferation. miR-21 was also one of the first miRNAs identified in the human genome that showed strong evolutionary conservation across a wide range

of vertebrate species. Three major targets of miR-21 include prominent tumor suppressors such as PTEN (phosphatase and tensin homolog), an important regulator of cardiovascular disease, PDCD4 (programmed cell death 4) and TPM1 (tropomyosin 1) [119–122].

The let-7 miRNA was one of the first anti-oncomiRs, or tumor suppressor miRNAs, characterized, which is highly conserved among mammalian species, and is downregulated in many tumors including lung and breast cancers [111,114,123]. The let-7 miRNA family functionally inhibits a number of well-characterized oncogenes such as *ras*, *c-myc* and *HMGA2* and induces apoptosis and cell cycle arrest in human colon cancer cells [123–126]. This miRNA targets the *ras* oncogene in lung cancer by being abnormally under-expressed promoting cell cycle progression [123]. In addition, let-7 also downregulates the expression of c-*myc*, a transcriptional activator of many tumor promoting genes that are dysregulated in lymphomas. Thus, anti-oncomiRs effectively control the expression of many oncogenes and their transcription factors at a post-transcriptional level.

4.2. Regulation of miRNAs in Human and Canine Breast Cancers

The association between altered miRNA expression signatures and breast cancer metastasis has been described by many studies [127,128]. A large number of miRNAs have been identified as deregulated in human breast cancer compared to normal breast tissue. The overexpression of certain oncogenic miRNAs (miR-21, miR-27a, miR-155, miR-9, miR-10b, miR-373/miR-520c, miR-206, miR-18a/b, miR-221/222) and the loss of several tumor suppressor miRNAs (miR-205/200, miR-125a, miR-125b, miR-126, miR-17-5p, miR-145, miR-200c, let-7, miR-20b, miR-34a, miR-31, miR-30) lead to loss of regulation of vital cellular functions that are involved in breast cancer pathogenesis [127,128]. In human breast cancer, miR-21 upregulates the EMT, the PI3K/ATK signaling pathway, the anti-apoptotic pathway and induces proliferation by targeting very well-characterized tumors suppressors such as PTEN, TPM1, and PDCD4 [121,122,129,130]. Strikingly, all of these miR-21 targets have been reported to be deregulated in canine mammary tumors as well. In this regard, expression of selected miRNAs associated with human breast cancers have been investigated in canine malignant mammary tumors. Almost all of the canine miRNAs in CMTs followed the same expression profile observed in human breast cancers when compared to normal canine mammary tissue. This investigation revealed that miR-21 and miR-29b were significantly up-regulated and miR-15b, miR-16 were significantly down-regulated in breast cancers in both species [131].

4.3. miRNAs Regulate Cell Cycle by Targeting Multiple Genes

An important function of miRNAs is to regulate cell cycle progression and arrest by targeting multiple cell cycle regulatory genes. These miRNAs regulate cell proliferation by specifically targeting cyclin-CDK complexes and CDK inhibitors. One of the first discoveries that connected miRNAs and cell cycle regulation was the anti-proliferative potential of the miR-15a/16-1 family that target multiple cell cycle genes involved in cellular proliferation and growth arrest [132–135]. The miR-16 family act as tumor suppressors that induce cell cycle arrest at the G1 phase by targeting several cyclin-CDK genes including CDK6, cyclin D1, cyclin D3, E2F3 and WEE1 and all the miRNAs in this family are downregulated in a wide variety of tumors [136]. Additionally, miR-34 and other family members, target CDK4/6, cyclin D1, cyclin E2, E2F1/3 and c-myc, indicating their strong anti-proliferative roles [137]. These miRNAs are transcriptionally activated by p53 and are involved in the p53 signaling pathway thereby acting as mediators of tumor growth suppression [138]. However, the tumor suppressive miRNAs involved in cell cycle regulation are inactivated in tumors by epigenetic mechanisms, such as hypermethylation, leading to overexpression of their target genes [139]. For example, members of the miR-290 family positively regulate G1 to S phase transition by inhibiting cyclin-dependent kinase inhibitors such as p21, during embryonic stem cell differentiation [140]. The Cip/Kip family CKIs are targeted by miR-17-92, miR-106b, the miR-221 family and miR-25 in many different carcinomas [136]. Expression of p16/INK4A is repressed by miR-24 and miR-31 which are also involved in the regulation of cell proliferation and progression

of cell cycle in many cancers [141,142]. It has been reported that miR-21 negatively regulates cell cycle during G1 to S phase transition in response to DNA damage and inhibits Cdc25A expression affecting G2/M progression in colon cancer cells [143]. Another study showed that miR-322/424 and miR-503 are upregulated during myogenesis and these miRNAs promote cell cycle arrest at G1 phase by down-regulating Cdc25A [144]. A recent report revealed that canine miR-141 can post-transcriptionally regulate p16/INK4A and p14ARF transcripts while groups of differentially expressed miRNAs may potentially target the rest of the CKI gene family members as well as oncogenes of the cell cycle in canine breast cancer models [145]. All of these reports clearly suggest that the cell cycle G1 to S phase transition is tightly regulated by several families of miRNAs. Therefore, different miRNAs regulate the cell cycle both positively and negatively by targeting the expression of many genes at different stages, and dysregulation of most of these regulatory molecules and pathways have been implicated in different pathological or developmental conditions.

5. Conclusions

In conclusion, the strong similarities in genome sequence, along with highly similar characteristics for spontaneous tumor models, have raised great promise for further comparative genomic research between humans and dogs. Comparing spontaneous mammary carcinomas in female dogs with breast cancer in women has significantly improved our understanding in deciphering the molecular mechanisms, relevant risk factors, and genetic profiles of these types of cancer and as well as novel strategies for future therapeutic inventions. However, although there is great potential in canine cancer models, a large number and complete interactions of cancer associated genes such as the cell cycle regulators, including the INK4 tumor suppressor genes and emerging miRNAs in the canine genome, have not been well studied in such models. Additionally, the high correlation between tumor suppressor gene expression and miRNA activity imposing post-transcriptional regulation is one of the central areas in cancer research which also needs to be further explored.

Acknowledgments: The authors (RCB and FMLK) thank AURIC (Auburn University Research Initiative in Cancer) for research and scholarship (FMLK) funding. CEA was supported by Department of Defense CDMRP award no. W81XWH-11-2-0224, American Kennel Club CHF award no. 01660, and Morris Animal Foundation award no. D13CA-073.

Author Contributions: Farruk M. Lutful Kabir and R. Curtis Bird worked together to develop the ideas presented in this review which has grown out of the doctoral training program of Luful Kabir. Specifically, Luful Kabir was a trainee in Bird's program developing these ideas as part of his training in full collaboration with Bird. Both authors contributed to the writing and editing of the manuscript. Carlos E. Alvarez provided in depth analysis of the genomes presented as well as the bioinformatics analysis and contributed to the editing of the manuscript as well.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rowell, J.L.; McCarthy, D.O.; Alvarez, C.E. Dog models of naturally occurring cancer. *Trends Mol. Med.* **2011**, *17*, 380–388. [CrossRef] [PubMed]
2. Smith, B.F.; Bird, R.C. Hematologic neoplasia—Gene therapy. In *Schalm's Veterinary Hematology*; Weiss, D.J., Wardrop, K.J., Eds.; Wiley-Blackwell: Iowa City, IA, USA, 2010; pp. 550–557.
3. American Veterinary Medical Association. *Us Pet Ownership and Demographics Sourcebook 2007*; American Veterinary Medical Association: Schaumburg, IL, USA, 2008.
4. Lindblad-Toh, K.; Wade, C.M.; Mikkelsen, T.S.; Karlsson, E.K.; Jaffe, D.B.; Kamal, M.; Clamp, M.; Chang, J.L.; Kulbokas, E.J., 3rd; Zody, M.C.; *et al*. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **2005**, *438*, 803–819. [CrossRef] [PubMed]
5. Khanna, C.; Lindblad-Toh, K.; Vail, D.; London, C.; Bergman, P.; Barber, L.; Breen, M.; Kitchell, B.; McNeil, E.; Modiano, J.F.; *et al*. The dog as a cancer model. *Nat. Biotechnol.* **2006**, *24*, 1065–1066. [CrossRef] [PubMed]

6. Vail, D.M.; MacEwen, E.G. Spontaneously occurring tumors of companion animals as models for human cancer. *Cancer Investig.* **2000**, *18*, 781–792. [CrossRef]
7. Cullen, J.M.; Page, R.; Misdorp, W. An overview of cancer pathogenesis, diagnosis and management. In *Tumors in Domestic Animals*; Moulton, D.J., Ed.; Blackwell Publishing Company, Iowa State Press: Ames, IA, USA, 2002; pp. 3–45.
8. Jemal, A.; Siegel, R.; Ward, E.; Murray, T.; Xu, J.; Thun, M.J. Cancer statistics, 2007. *CA* **2007**, *57*, 43–66. [CrossRef] [PubMed]
9. Owen, L.N. A comparative study of canine and human breast cancer. *Investig. Cell Pathol.* **1979**, *2*, 257–275.
10. Sorenmo, K. Canine mammary gland tumors. *Vet. Clin. North Am. Small Anim. Pract.* **2003**, *33*, 573–596. [CrossRef]
11. Ahern, T.E.; Bird, R.C.; Bird, A.E.; Wolfe, L.G. Expression of the oncogene c-erbb-2 in canine mammary cancers and tumor-derived cell lines. *Am. J. Vet. Res.* **1996**, *57*, 693–696. [PubMed]
12. Misdorp, W. Tumors of the mammary gland. In *Tumors in Domestic Animals*; Meuten, D.J., Ed.; Iowa State Press: Ames, IA, USA, 2002; pp. 575–606.
13. Sleeckx, N.; de Rooster, H.; Veldhuis Kroeze, E.J.; Van Ginneken, C.; Van Brantegem, L. Canine mammary tumours, an overview. *Reprod. Domest. Anim.* **2011**, *46*, 1112–1131. [CrossRef] [PubMed]
14. Shinoda, H.; Legare, M.E.; Mason, G.L.; Berkbigler, J.L.; Afzali, M.F.; Flint, A.F.; Hanneman, W.H. Significance of eralpha, her2, and cav1 expression and molecular subtype classification to canine mammary gland tumor. *J. Vet. Diagn. Investig.* **2014**, *26*, 390–403. [CrossRef] [PubMed]
15. Gama, A.; Alves, A.; Schmitt, F. Identification of molecular phenotypes in canine mammary carcinomas with clinical implications: Application of the human classification. *Virchows Arch.* **2008**, *453*, 123–132. [CrossRef] [PubMed]
16. Beck, J.; Hennecke, S.; Bornemann-Kolatzki, K.; Urnovitz, H.B.; Neumann, S.; Strobel, P.; Kaup, F.J.; Brenig, B.; Schutz, E. Genome aberrations in canine mammary carcinomas and their detection in cell-free plasma DNA. *PLoS ONE* **2013**, *8*, e75485. [CrossRef] [PubMed]
17. Sassi, F.; Benazzi, C.; Castellani, G.; Sarli, G. Molecular-based tumour subtypes of canine mammary carcinomas assessed by immunohistochemistry. *BMC Vet. Res.* **2010**, *6*, 5. [CrossRef] [PubMed]
18. Lutful Kabir, F.M.; DeInnocentes, P.; Agarwal, P.; Riese, D.J.; Bird, R.C. Estrogen receptor-alpha, progesterone receptor and c-erbb/her-family receptor mrna detection and phenotype analysis in spontaneous canine models of breast cancer. *Vet. Pathol.* **2015**, in preparation.
19. Uva, P.; Aurisicchio, L.; Watters, J.; Loboda, A.; Kulkarni, A.; Castle, J.; Palombo, F.; Viti, V.; Mesiti, G.; Zappulli, V.; *et al*. Comparative expression pathway analysis of human and canine mammary tumors. *BMC Genom.* **2009**, *10*, 135. [CrossRef] [PubMed]
20. Rao, N.A.; van Wolferen, M.E.; van den Ham, R.; van Leenen, D.; Groot Koerkamp, M.J.; Holstege, F.C.; Mol, J.A. Cdna microarray profiles of canine mammary tumour cell lines reveal deregulated pathways pertaining to their phenotype. *Anim. Genet.* **2008**, *39*, 333–345. [CrossRef] [PubMed]
21. Lutful Kabir, F.M.; Agarwal, P.; Deinnocentes, P.; Zaman, J.; Bird, A.C.; Bird, R.C. Novel frameshift mutation in the p16/ink4a tumor suppressor gene in canine breast cancer alters expression from the p16/ink4a/p14arf locus. *J. Cell. Biochem.* **2013**, *114*, 56–66. [CrossRef] [PubMed]
22. Klopfleisch, R.; Gruber, A.D. Differential expression of cell cycle regulators p21, p27 and p53 in metastasizing canine mammary adenocarcinomas *versus* normal mammary glands. *Res. Vet. Sci.* **2009**, *87*, 91–96. [CrossRef] [PubMed]
23. Van 't Veer, L.J.; Dai, H.; van de Vijver, M.J.; He, Y.D.; Hart, A.A.; Mao, M.; Peterse, H.L.; van der Kooy, K.; Marton, M.J.; Witteveen, A.T.; *et al*. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **2002**, *415*, 530–536. [CrossRef] [PubMed]
24. Saal, L.H.; Johansson, P.; Holm, K.; Gruvberger-Saal, S.K.; She, Q.B.; Maurer, M.; Koujak, S.; Ferrando, A.A.; Malmstrom, P.; Memeo, L.; *et al*. Poor prognosis in carcinoma is associated with a gene expression signature of aberrant pten tumor suppressor pathway activity. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7564–7569. [CrossRef] [PubMed]
25. Sweet-Cordero, A.; Mukherjee, S.; Subramanian, A.; You, H.; Roix, J.J.; Ladd-Acosta, C.; Mesirov, J.; Golub, T.R.; Jacks, T. An oncogenic kras2 expression signature identified by cross-species gene-expression analysis. *Nat. Genet.* **2005**, *37*, 48–55. [CrossRef] [PubMed]

26. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [CrossRef] [PubMed]
27. Pines, J.; Hunter, T. Cyclin-dependent kinases: A new cell cycle motif? *Trends Cell Biol.* **1991**, *1*, 117–121. [CrossRef]
28. Mareel, M.; Leroy, A. Clinical, cellular, and molecular aspects of cancer invasion. *Physiol. Rev.* **2003**, *83*, 337–376. [CrossRef] [PubMed]
29. Sherr, C.J. Cancer cell cycles. *Science* **1996**, *274*, 1672–1677. [CrossRef] [PubMed]
30. Okamoto, A.; Demetrick, D.J.; Spillare, E.A.; Hagiwara, K.; Hussain, S.P.; Bennett, W.P.; Forrester, K.; Gerwin, B.; Serrano, M.; Beach, D.H.; *et al*. Mutations and altered expression of p16ink4 in human cancer. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11045–11049. [CrossRef] [PubMed]
31. Otterson, G.A.; Kratzke, R.A.; Coxon, A.; Kim, Y.W.; Kaye, F.J. Absence of p16ink4 protein is restricted to the subset of lung cancer lines that retains wildtype rb. *Oncogene* **1994**, *9*, 3375–3378. [PubMed]
32. Harper, J.W.; Elledge, S.J. Cdk inhibitors in development and cancer. *Curr. Opin. Genet. Dev.* **1996**, *6*, 56–64. [CrossRef]
33. Vidal, A.; Koff, A. Cell-cycle inhibitors: Three families united by a common cause. *Gene* **2000**, *247*, 1–15. [CrossRef]
34. McConnell, B.B.; Gregory, F.J.; Stott, F.J.; Hara, E.; Peters, G. Induced expression of p16(ink4a) inhibits both cdk4- and cdk2-associated kinase activity by reassortment of cyclin-cdk-inhibitor complexes. *Mol. Cell. Biol.* **1999**, *19*, 1981–1989. [CrossRef] [PubMed]
35. Parry, D.; Mahony, D.; Wills, K.; Lees, E. Cyclin d-cdk subunit arrangement is dependent on the availability of competing ink4 and p21 class inhibitors. *Mol. Cell. Biol.* **1999**, *19*, 1775–1783. [CrossRef] [PubMed]
36. Ruas, M.; Peters, G. The p16ink4a/cdkn2a tumor suppressor and its relatives. *Biochim. Biophys. Acta* **1998**, *1378*, F115–F177. [CrossRef]
37. Sharpless, N.E. Ink4a/arf: A multifunctional tumor suppressor locus. *Mutat. Res.* **2005**, *576*, 22–38. [CrossRef] [PubMed]
38. Sharpless, N.E.; DePinho, R.A. The ink4a/arf locus and its two gene products. *Curr. Opin. Genet. Dev.* **1999**, *9*, 22–30. [CrossRef]
39. Xiong, Y.; Hannon, G.J.; Zhang, H.; Casso, D.; Kobayashi, R.; Beach, D. P21 is a universal inhibitor of cyclin kinases. *Nature* **1993**, *366*, 701–704. [CrossRef] [PubMed]
40. Hunter, T.; Pines, J. Cyclins and cancer. Ii: Cyclin d and cdk inhibitors come of age. *Cell* **1994**, *79*, 573–582. [CrossRef]
41. Kamb, A.; Gruis, N.A.; Weaver-Feldhaus, J.; Liu, Q.; Harshman, K.; Tavtigian, S.V.; Stockert, E.; Day, R.S., 3rd; Johnson, B.E.; Skolnick, M.H. A cell cycle regulator potentially involved in genesis of many tumor types. *Science* **1994**, *264*, 436–440. [CrossRef] [PubMed]
42. Wesierska-Gadek, J.; Schmid, G. Dual action of the inhibitors of cyclin-dependent kinases: Targeting of the cell-cycle progression and activation of wild-type p53 protein. *Expert Opin. Investig. Drugs* **2006**, *15*, 23–38. [CrossRef] [PubMed]
43. Serrano, M. The ink4a/arf locus in murine tumorigenesis. *Carcinogenesis* **2000**, *21*, 865–869. [CrossRef] [PubMed]
44. Aguirre-Hernandez, J.; Milne, B.S.; Queen, C.; O'Brien, P.C.; Hoather, T.; Haugland, S.; Ferguson-Smith, M.A.; Dobson, J.M.; Sargan, D.R. Disruption of chromosome 11 in canine fibrosarcomas highlights an unusual variability of cdkn2b in dogs. *BMC Vet. Res.* **2009**, *5*, 27. [CrossRef] [PubMed]
45. DeInnocentes, P.; Agarwal, P.; Bird, R.C. Phenotype-rescue of cyclin-dependent kinase inhibitor p16/ink4a defects in a spontaneous canine cell model of breast cancer. *J. Cell. Biochem.* **2009**, *106*, 491–505. [CrossRef] [PubMed]
46. Koenig, A.; Bianco, S.R.; Fosmire, S.; Wojcieszyn, J.; Modiano, J.F. Expression and significance of p53, rb, p21/waf-1, p16/ink-4a, and pten tumor suppressors in canine melanoma. *Vet. Pathol.* **2002**, *39*, 458–472. [CrossRef] [PubMed]
47. Migone, F.; Deinnocentes, P.; Smith, B.F.; Bird, R.C. Alterations in cdk1 expression and nuclear/nucleolar localization following induction in a spontaneous canine mammary cancer model. *J. Cell. Biochem.* **2006**, *98*, 504–518. [CrossRef] [PubMed]
48. Gilley, J.; Fried, M. One ink4 gene and no arf at the fugu equivalent of the human ink4a/arf/ink4b tumour suppressor locus. *Oncogene* **2001**, *20*, 7447–7452. [CrossRef] [PubMed]

49. Byeon, I.J.; Li, J.; Ericson, K.; Selby, T.L.; Tevelev, A.; Kim, H.J.; O'Maille, P.; Tsai, M.D. Tumor suppressor p16ink4a: Determination of solution structure and analyses of its interaction with cyclin-dependent kinase 4. *Mol. Cell* **1998**, *1*, 421–431. [CrossRef]
50. Venkataramani, R.; Swaminathan, K.; Marmorstein, R. Crystal structure of the cdk4/6 inhibitory protein p18ink4c provides insights into ankyrin-like repeat structure/function and tumor-derived p16ink4 mutations. *Nat. Struct. Biol.* **1998**, *5*, 74–81. [CrossRef] [PubMed]
51. Mao, L.; Merlo, A.; Bedi, G.; Shapiro, G.I.; Edwards, C.D.; Rollins, B.J.; Sidransky, D. A novel p16ink4a transcript. *Cancer Res.* **1995**, *55*, 2995–2997. [PubMed]
52. Guan, K.L.; Jenkins, C.W.; Li, Y.; O'Keefe, C.L.; Noh, S.; Wu, X.; Zariwala, M.; Matera, A.G.; Xiong, Y. Isolation and characterization of p19ink4d, a p16-related inhibitor specific to cdk6 and cdk4. *Mol. Biol. Cell* **1996**, *7*, 57–70. [CrossRef] [PubMed]
53. Sherr, C.J. G1 phase progression: Cycling on cue. *Cell* **1994**, *79*, 551–555. [CrossRef]
54. Weinberg, R.A. The retinoblastoma protein and cell cycle control. *Cell* **1995**, *81*, 323–330. [CrossRef]
55. Quelle, D.E.; Ashmun, R.A.; Hannon, G.J.; Rehberger, P.A.; Trono, D.; Richter, K.H.; Walker, C.; Beach, D.; Sherr, C.J.; Serrano, M. Cloning and characterization of murine p16ink4a and p15ink4b genes. *Oncogene* **1995**, *11*, 635–645. [PubMed]
56. Chin, L.; Pomerantz, J.; Polsky, D.; Jacobson, M.; Cohen, C.; Cordon-Cardo, C.; Horner, J.W., 2nd; DePinho, R.A. Cooperative effects of ink4a and ras in melanoma susceptibility *in vivo*. *Genes Dev.* **1997**, *11*, 2822–2834. [CrossRef] [PubMed]
57. Kamijo, T.; Zindy, F.; Roussel, M.F.; Quelle, D.E.; Downing, J.R.; Ashmun, R.A.; Grosveld, G.; Sherr, C.J. Tumor suppression at the mouse ink4a locus mediated by the alternative reading frame product p19arf. *Cell* **1997**, *91*, 649–659. [CrossRef]
58. Pomerantz, J.; Schreiber-Agus, N.; Liegeois, N.J.; Silverman, A.; Alland, L.; Chin, L.; Potes, J.; Chen, K.; Orlow, I.; Lee, H.W.; *et al*. The ink4a tumor suppressor gene product, p19arf, interacts with mdm2 and neutralizes mdm2's inhibition of p53. *Cell* **1998**, *92*, 713–723. [CrossRef]
59. Kamijo, T.; Weber, J.D.; Zambetti, G.; Zindy, F.; Roussel, M.F.; Sherr, C.J. Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8292–8297. [CrossRef] [PubMed]
60. Hussussian, C.J.; Struewing, J.P.; Goldstein, A.M.; Higgins, P.A.; Ally, D.S.; Sheahan, M.D.; Clark, W.H., Jr.; Tucker, M.A.; Dracopoli, N.C. Germline p16 mutations in familial melanoma. *Nat. Genet.* **1994**, *8*, 15–21. [CrossRef] [PubMed]
61. Kamb, A.; Shattuck-Eidens, D.; Eeles, R.; Liu, Q.; Gruis, N.A.; Ding, W.; Hussey, C.; Tran, T.; Miki, Y.; Weaver-Feldhaus, J.; *et al*. Analysis of the p16 gene (cdkn2) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nat. Genet.* **1994**, *8*, 23–26. [CrossRef] [PubMed]
62. Ranade, K.; Hussussian, C.J.; Sikorski, R.S.; Varmus, H.E.; Goldstein, A.M.; Tucker, M.A.; Serrano, M.; Hannon, G.J.; Beach, D.; Dracopoli, N.C. Mutations associated with familial melanoma impair p16ink4 function. *Nat. Genet.* **1995**, *10*, 114–116. [CrossRef] [PubMed]
63. Caldas, C.; Hahn, S.A.; da Costa, L.T.; Redston, M.S.; Schutte, M.; Seymour, A.B.; Weinstein, C.L.; Hruban, R.H.; Yeo, C.J.; Kern, S.E. Frequent somatic mutations and homozygous deletions of the p16 (mts1) gene in pancreatic adenocarcinoma. *Nat. Genet.* **1994**, *8*, 27–32. [CrossRef] [PubMed]
64. Dreyling, M.H.; Bohlander, S.K.; Adeyanju, M.O.; Olopade, O.I. Detection of cdkn2 deletions in tumor cell lines and primary glioma by interphase fluorescence *in situ* hybridization. *Cancer Res.* **1995**, *55*, 984–988. [PubMed]
65. Hatta, Y.; Hirama, T.; Miller, C.W.; Yamada, Y.; Tomonaga, M.; Koeffler, H.P. Homozygous deletions of the p15 (mts2) and p16 (cdkn2/mts1) genes in adult t-cell leukemia. *Blood* **1995**, *85*, 2699–2704. [PubMed]
66. Nobori, T.; Miura, K.; Wu, D.J.; Lois, A.; Takabayashi, K.; Carson, D.A. Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. *Nature* **1994**, *368*, 753–756. [CrossRef] [PubMed]
67. Ogawa, S.; Hirano, N.; Sato, N.; Takahashi, T.; Hangaishi, A.; Tanaka, K.; Kurokawa, M.; Tanaka, T.; Mitani, K.; Yazaki, Y.; *et al*. Homozygous loss of the cyclin-dependent kinase 4-inhibitor (p16) gene in human leukemias. *Blood* **1994**, *84*, 2431–2435. [PubMed]
68. Quesnel, B.; Preudhomme, C.; Philippe, N.; Vanrumbeke, M.; Dervite, I.; Lai, J.L.; Bauters, F.; Wattel, E.; Fenaux, P. P16 gene homozygous deletions in acute lymphoblastic leukemia. *Blood* **1995**, *85*, 657–663. [PubMed]

69. Carrera, C.J.; Eddy, R.L.; Shows, T.B.; Carson, D.A. Assignment of the gene for methylthioadenosine phosphorylase to human chromosome 9 by mouse-human somatic cell hybridization. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 2665–2668. [CrossRef] [PubMed]
70. Nobori, T.; Takabayashi, K.; Tran, P.; Orvis, L.; Batova, A.; Yu, A.L.; Carson, D.A. Genomic cloning of methylthioadenosine phosphorylase: A purine metabolic enzyme deficient in multiple different cancers. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6203–6208. [CrossRef] [PubMed]
71. Herzog, C.R.; Noh, S.; Lantry, L.E.; Guan, K.L.; You, M. Cdkn2a encodes functional variation of p16ink4a but not p19arf, which confers selection in mouse lung tumorigenesis. *Mol. Carcinog.* **1999**, *25*, 92–98. [CrossRef]
72. Zhang, S.; Ramsay, E.S.; Mock, B.A. Cdkn2a, the cyclin-dependent kinase inhibitor encoding p16ink4a and p19arf, is a candidate for the plasmacytoma susceptibility locus, pctr1. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2429–2434. [CrossRef] [PubMed]
73. Krimpenfort, P.; Quon, K.C.; Mooi, W.J.; Loonstra, A.; Berns, A. Loss of p16ink4a confers susceptibility to metastatic melanoma in mice. *Nature* **2001**, *413*, 83–86. [CrossRef] [PubMed]
74. Serrano, M.; Lee, H.; Chin, L.; Cordon-Cardo, C.; Beach, D.; DePinho, R.A. Role of the ink4a locus in tumor suppression and cell mortality. *Cell* **1996**, *85*, 27–37. [CrossRef]
75. Sharpless, N.E.; Bardeesy, N.; Lee, K.H.; Carrasco, D.; Castrillon, D.H.; Aguirre, A.J.; Wu, E.A.; Horner, J.W.; DePinho, R.A. Loss of p16ink4a with retention of p19arf predisposes mice to tumorigenesis. *Nature* **2001**, *413*, 86–91. [CrossRef] [PubMed]
76. Levine, R.A.; Fleischli, M.A. Inactivation of p53 and retinoblastoma family pathways in canine osteosarcoma cell lines. *Vet. Pathol.* **2000**, *37*, 54–61. [CrossRef] [PubMed]
77. Yeudall, W.A.; Crawford, R.Y.; Ensley, J.F.; Robbins, K.C. Mts1/cdk4i is altered in cell lines derived from primary and metastatic oral squamous cell carcinoma. *Carcinogenesis* **1994**, *15*, 2683–2686. [CrossRef] [PubMed]
78. Zhang, S.L.; DuBois, W.; Ramsay, E.S.; Bliskovski, V.; Morse, H.C., 3rd; Taddesse-Heath, L.; Vass, W.C.; DePinho, R.A.; Mock, B.A. Efficiency alleles of the pctr1 modifier locus for plasmacytoma susceptibility. *Mol. Cell. Biol.* **2001**, *21*, 310–318. [CrossRef] [PubMed]
79. Shearin, A.L.; Hedan, B.; Cadieu, E.; Erich, S.A.; Schmidt, E.V.; Faden, D.L.; Cullen, J.; Abadie, J.; Kwon, E.M.; Grone, A.; *et al*. The mtap-cdkn2a locus confers susceptibility to a naturally occurring canine cancer. *Cancer Epidemiol. Biomark. Prev.* **2012**, *21*, 1019–1027. [CrossRef] [PubMed]
80. Melkamu, T.; Zhang, X.; Tan, J.; Zeng, Y.; Kassie, F. Alteration of microrna expression in vinyl carbamate-induced mouse lung tumors and modulation by the chemopreventive agent indole-3-carbinol. *Carcinogenesis* **2010**, *31*, 252–258. [CrossRef] [PubMed]
81. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigo, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, E.T.; Thurman, R.E.; *et al*. Identification and analysis of functional elements in 1% of the human genome by the encode pilot project. *Nature* **2007**, *447*, 799–816. [CrossRef] [PubMed]
82. Costa, F.F. Non-coding rnas: Meet thy masters. *BioEssays: News Rev. Mol. Cell. Dev. Biol.* **2010**, *32*, 599–608. [CrossRef] [PubMed]
83. Prensner, J.R.; Chinnaiyan, A.M. The emergence of lncrnas in cancer biology. *Cancer Discov.* **2011**, *1*, 391–407. [CrossRef] [PubMed]
84. Cheetham, S.W.; Gruhl, F.; Mattick, J.S.; Dinger, M.E. Long noncoding rnas and the genetics of cancer. *Br. J. Cancer* **2013**, *108*, 2419–2425. [CrossRef] [PubMed]
85. Gibb, E.A.; Brown, C.J.; Lam, W.L. The functional role of long non-coding rna in human carcinomas. *Mol. Cancer* **2011**, *10*, 38. [CrossRef] [PubMed]
86. Mercer, T.R.; Dinger, M.E.; Sunkin, S.M.; Mehler, M.F.; Mattick, J.S. Specific expression of long noncoding rnas in the mouse brain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 716–721. [CrossRef] [PubMed]
87. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al*. Long non-coding rna hotair reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076. [CrossRef] [PubMed]
88. Morey, L.; Helin, K. Polycomb group protein-mediated repression of transcription. *Trends Biochem. Sci.* **2010**, *35*, 323–332. [CrossRef] [PubMed]

89. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding rna as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693. [CrossRef] [PubMed]
90. Zhang, Z.; Jones, A.; Sun, C.W.; Li, C.; Chang, C.W.; Joo, H.Y.; Dai, Q.; Mysliwiec, M.R.; Wu, L.C.; Guo, Y.; *et al*. Prc2 complexes with jarid2, mtf2, and esprc2p48 in es cells to modulate es cell pluripotency and somatic cell reprogramming. *Stem Cells* **2011**, *29*, 229–240. [CrossRef] [PubMed]
91. Jovanovic, M.; Hengartner, M.O. Mirnas and apoptosis: Rnas to die for. *Oncogene* **2006**, *25*, 6176–6187. [CrossRef] [PubMed]
92. Schickel, R.; Boyerinas, B.; Park, S.M.; Peter, M.E. Micrornas: Key players in the immune system, differentiation, tumorigenesis and cell death. *Oncogene* **2008**, *27*, 5959–5974. [CrossRef] [PubMed]
93. Stefani, G.; Slack, F.J. Small non-coding rnas in animal development. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 219–230. [CrossRef] [PubMed]
94. Lagos-Quintana, M.; Rauhut, R.; Lendeckel, W.; Tuschl, T. Identification of novel genes coding for small expressed rnas. *Science* **2001**, *294*, 853–858. [CrossRef] [PubMed]
95. Esquela-Kerscher, A.; Slack, F.J. Oncomirs—Micrornas with a role in cancer. *Nat. Rev. Cancer* **2006**, *6*, 259–269. [CrossRef] [PubMed]
96. Calin, G.A.; Sevignani, C.; Dumitru, C.D.; Hyslop, T.; Noch, E.; Yendamuri, S.; Shimizu, M.; Rattan, S.; Bullrich, F.; Negrini, M.; *et al*. Human microrna genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2999–3004. [CrossRef] [PubMed]
97. Garzon, R.; Calin, G.A.; Croce, C.M. Micrornas in cancer. *Annu. Rev. Med.* **2009**, *60*, 167–179. [CrossRef] [PubMed]
98. Bartel, D.P. Micrornas: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297. [CrossRef]
99. Fabian, M.R.; Sonenberg, N.; Filipowicz, W. Regulation of mrna translation and stability by micrornas. *Annu. Rev. Biochem.* **2010**, *79*, 351–379. [CrossRef] [PubMed]
100. Friedman, R.C.; Farh, K.K.; Burge, C.B.; Bartel, D.P. Most mammalian mrnas are conserved targets of micrornas. *Genome Res.* **2009**, *19*, 92–105. [CrossRef] [PubMed]
101. Kent, O.A.; Mendell, J.T. A small piece in the cancer puzzle: Micrornas as tumor suppressors and oncogenes. *Oncogene* **2006**, *25*, 6188–6196. [CrossRef] [PubMed]
102. Hata, A.; Kashima, R. Dysregulation of microRNA biogenesis machinery in cancer. *Crit. Rev. Biochem. Molec. Biol.* **2015**. [CrossRef] [PubMed]
103. Chendrimada, T.P.; Gregory, R.I.; Kumaraswamy, E.; Norman, J.; Cooch, N.; Nishikura, K.; Shiekhattar, R. Trbp recruits the dicer complex to ago2 for microrna processing and gene silencing. *Nature* **2005**, *436*, 740–744. [CrossRef] [PubMed]
104. Hammond, S.M.; Bernstein, E.; Beach, D.; Hannon, G.J. An rna-directed nuclease mediates post-transcriptional gene silencing in drosophila cells. *Nature* **2000**, *404*, 293–296. [PubMed]
105. Zeng, Y.; Wagner, E.J.; Cullen, B.R. Both natural and designed micro rnas can inhibit the expression of cognate mrnas when expressed in human cells. *Mol. Cell* **2002**, *9*, 1327–1333. [CrossRef]
106. Martello, G.; Rosato, A.; Ferrari, F.; Manfrin, A.; Cordenonsi, M.; Dupont, S.; Enzo, E.; Guzzardo, V.; Rondina, M.; Spruce, T.; *et al*. A microrna targeting dicer for metastasis control. *Cell* **2010**, *141*, 1195–1207. [CrossRef] [PubMed]
107. Calin, G.A.; Dumitru, C.D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; *et al*. Frequent deletions and down-regulation of micro- rna genes mir15 and mir16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15524–15529. [CrossRef] [PubMed]
108. Chang, T.C.; Wentzel, E.A.; Kent, O.A.; Ramachandran, K.; Mullendore, M.; Lee, K.H.; Feldmann, G.; Yamakuchi, M.; Ferlito, M.; Lowenstein, C.J.; *et al*. Transactivation of mir-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol. Cell* **2007**, *26*, 745–752. [CrossRef] [PubMed]
109. Saito, Y.; Liang, G.; Egger, G.; Friedman, J.M.; Chuang, J.C.; Coetzee, G.A.; Jones, P.A. Specific activation of microrna-127 with downregulation of the proto-oncogene bcl6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **2006**, *9*, 435–443. [CrossRef] [PubMed]
110. Calin, G.A.; Ferracin, M.; Cimmino, A.; Di Leva, G.; Shimizu, M.; Wojcik, S.E.; Iorio, M.V.; Visone, R.; Sever, N.I.; Fabbri, M.; *et al*. A microrna signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.* **2005**, *353*, 1793–1801. [CrossRef] [PubMed]

111. Iorio, M.V.; Ferracin, M.; Liu, C.G.; Veronese, A.; Spizzo, R.; Sabbioni, S.; Magri, E.; Pedriali, M.; Fabbri, M.; Campiglio, M.; *et al*. Microrna gene expression deregulation in human breast cancer. *Cancer Res.* **2005**, *65*, 7065–7070. [CrossRef] [PubMed]
112. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; *et al*. Microrna expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838. [CrossRef] [PubMed]
113. Volinia, S.; Calin, G.A.; Liu, C.G.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M.; *et al*. A microrna expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2257–2261. [CrossRef] [PubMed]
114. Yanaihara, N.; Caplen, N.; Bowman, E.; Seike, M.; Kumamoto, K.; Yi, M.; Stephens, R.M.; Okamoto, A.; Yokota, J.; Tanaka, T.; *et al*. Unique microrna molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* **2006**, *9*, 189–198. [CrossRef] [PubMed]
115. Kluiver, J.; Poppema, S.; de Jong, D.; Blokzijl, T.; Harms, G.; Jacobs, S.; Kroesen, B.J.; van den Berg, A. Bic and mir-155 are highly expressed in hodgkin, primary mediastinal and diffuse large b cell lymphomas. *J. Pathol.* **2005**, *207*, 243–249. [CrossRef] [PubMed]
116. Metzler, M.; Wilda, M.; Busch, K.; Viehmann, S.; Borkhardt, A. High expression of precursor microrna-155/bic rna in children with burkitt lymphoma. *Genes Chromosomes Cancer* **2004**, *39*, 167–169. [CrossRef] [PubMed]
117. Ciafre, S.A.; Galardi, S.; Mangiola, A.; Ferracin, M.; Liu, C.G.; Sabatino, G.; Negrini, M.; Maira, G.; Croce, C.M.; Farace, M.G. Extensive modulation of a set of micrornas in primary glioblastoma. *Biochem. Biophys. Res. Commun.* **2005**, *334*, 1351–1358. [CrossRef] [PubMed]
118. Garzon, R.; Garofalo, M.; Martelli, M.P.; Briesewitz, R.; Wang, L.; Fernandez-Cymering, C.; Volinia, S.; Liu, C.G.; Schnittger, S.; Haferlach, T.; *et al*. Distinctive microrna signature of acute myeloid leukemia bearing cytoplasmic mutated nucleophosmin. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3945–3950. [CrossRef] [PubMed]
119. Meng, F.; Henson, R.; Wehbe-Janek, H.; Ghoshal, K.; Jacob, S.T.; Patel, T. Microrna-21 regulates expression of the pten tumor suppressor gene in human hepatocellular cancer. *Gastroenterology* **2007**, *133*, 647–658. [CrossRef] [PubMed]
120. Chan, J.A.; Krichevsky, A.M.; Kosik, K.S. Microrna-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res.* **2005**, *65*, 6029–6033. [CrossRef] [PubMed]
121. Frankel, L.B.; Christoffersen, N.R.; Jacobsen, A.; Lindow, M.; Krogh, A.; Lund, A.H. Programmed cell death 4 (pdcd4) is an important functional target of the microrna mir-21 in breast cancer cells. *J. Biol. Chem.* **2008**, *283*, 1026–1033. [CrossRef] [PubMed]
122. Zhu, S.; Si, M.L.; Wu, H.; Mo, Y.Y. Microrna-21 targets the tumor suppressor gene tropomyosin 1 (tpm1). *J. Biol. Chem.* **2007**, *282*, 14328–14336. [CrossRef] [PubMed]
123. Johnson, S.M.; Grosshans, H.; Shingara, J.; Byrom, M.; Jarvis, R.; Cheng, A.; Labourier, E.; Reinert, K.L.; Brown, D.; Slack, F.J. Ras is regulated by the let-7 microrna family. *Cell* **2005**, *120*, 635–647. [CrossRef] [PubMed]
124. Akao, Y.; Nakagawa, Y.; Naoe, T. Let-7 microrna functions as a potential growth suppressor in human colon cancer cells. *Biol. Pharm. Bull.* **2006**, *29*, 903–906. [CrossRef] [PubMed]
125. Lee, Y.S.; Dutta, A. The tumor suppressor microrna let-7 represses the hmga2 oncogene. *Genes Dev.* **2007**, *21*, 1025–1030. [CrossRef] [PubMed]
126. Sampson, V.B.; Rong, N.H.; Han, J.; Yang, Q.; Aris, V.; Soteropoulos, P.; Petrelli, N.J.; Dunn, S.P.; Krueger, L.J. Microrna let-7a down-regulates myc and reverts myc-induced growth in burkitt lymphoma cells. *Cancer Res.* **2007**, *67*, 9762–9770. [CrossRef] [PubMed]
127. Harquail, J.; Benzina, S.; Robichaud, G.A. Micrornas and breast cancer malignancy: An overview of mirna-regulated cancer processes leading to metastasis. *Cancer Biomark.: Sect. Dis. Markers* **2012**, *11*, 269–280.
128. Zhang, Z.J.; Ma, S.L. Mirnas in breast cancer tumorigenesis (review). *Oncol. Rep.* **2012**, *27*, 903–910. [PubMed]
129. Qi, L.; Bart, J.; Tan, L.P.; Platteel, I.; Sluis, T.; Huitema, S.; Harms, G.; Fu, L.; Hollema, H.; Berg, A. Expression of mir-21 and its targets (pten, pdcd4, tm1) in flat epithelial atypia of the breast in relation to ductal carcinoma *in situ* and invasive carcinoma. *BMC Cancer* **2009**, *9*, 163. [CrossRef] [PubMed]

130. Zhu, S.; Wu, H.; Wu, F.; Nie, D.; Sheng, S.; Mo, Y.Y. Microrna-21 targets tumor suppressor genes in invasion and metastasis. *Cell Res.* **2008**, *18*, 350–359. [CrossRef] [PubMed]
131. Boggs, R.M.; Wright, Z.M.; Stickney, M.J.; Porter, W.W.; Murphy, K.E. Microrna expression in canine mammary cancer. *Mamm. Genome* **2008**, *19*, 561–569. [CrossRef] [PubMed]
132. Linsley, P.S.; Schelter, J.; Burchard, J.; Kibukawa, M.; Martin, M.M.; Bartz, S.R.; Johnson, J.M.; Cummins, J.M.; Raymond, C.K.; Dai, H.; *et al*. Transcripts targeted by the microrna-16 family cooperatively regulate cell cycle progression. *Mol. Cell. Biol.* **2007**, *27*, 2240–2252. [CrossRef] [PubMed]
133. Liu, Q.; Fu, H.; Sun, F.; Zhang, H.; Tie, Y.; Zhu, J.; Xing, R.; Sun, Z.; Zheng, X. Mir-16 family induces cell cycle arrest by regulating multiple cell cycle genes. *Nucleic Acids Res.* **2008**, *36*, 5391–5404. [CrossRef] [PubMed]
134. Takeshita, F.; Patrawala, L.; Osaki, M.; Takahashi, R.U.; Yamamoto, Y.; Kosaka, N.; Kawamata, M.; Kelnar, K.; Bader, A.G.; Brown, D.; *et al*. Systemic delivery of synthetic microrna-16 inhibits the growth of metastatic prostate tumors via downregulation of multiple cell-cycle genes. *Mol. Ther.* **2010**, *18*, 181–187. [CrossRef] [PubMed]
135. Wang, F.; Fu, X.D.; Zhou, Y.; Zhang, Y. Down-regulation of the cyclin e1 oncogene expression by microrna-16–1 induces cell cycle arrest in human cancer cells. *BMB Rep.* **2009**, *42*, 725–730. [CrossRef] [PubMed]
136. Bueno, M.J.; Malumbres, M. Micrornas and the cell cycle. *Biochim. Biophys. Acta* **2011**, *1812*, 592–601. [CrossRef] [PubMed]
137. Sun, F.; Fu, H.; Liu, Q.; Tie, Y.; Zhu, J.; Xing, R.; Sun, Z.; Zheng, X. Downregulation of ccnd1 and cdk6 by mir-34a induces cell cycle arrest. *FEBS Lett.* **2008**, *582*, 1564–1568. [CrossRef] [PubMed]
138. He, L.; He, X.; Lim, L.P.; de Stanchina, E.; Xuan, Z.; Liang, Y.; Xue, W.; Zender, L.; Magnus, J.; Ridzon, D.; *et al*. A microrna component of the p53 tumour suppressor network. *Nature* **2007**, *447*, 1130–1134. [CrossRef] [PubMed]
139. Lujambio, A.; Ropero, S.; Ballestar, E.; Fraga, M.F.; Cerrato, C.; Setien, F.; Casado, S.; Suarez-Gauthier, A.; Sanchez-Cespedes, M.; Git, A.; *et al*. Genetic unmasking of an epigenetically silenced microrna in human cancer cells. *Cancer Res.* **2007**, *67*, 1424–1429. [CrossRef] [PubMed]
140. Wang, Y.; Blelloch, R. Cell cycle regulation by micrornas in embryonic stem cells. *Cancer Res.* **2009**, *69*, 4093–4096. [CrossRef] [PubMed]
141. Lal, A.; Kim, H.H.; Abdelmohsen, K.; Kuwano, Y.; Pullmann, R., Jr.; Srikantan, S.; Subrahmanyam, R.; Martindale, J.L.; Yang, X.; Ahmed, F.; *et al*. P16(ink4a) translation suppressed by mir-24. *PLoS ONE* **2008**, *3*, e1864. [CrossRef] [PubMed]
142. Malhas, A.; Saunders, N.J.; Vaux, D.J. The nuclear envelope can control gene expression and cell cycle progression via mirna regulation. *Cell Cycle* **2010**, *9*, 531–539. [CrossRef] [PubMed]
143. Wang, P.; Zou, F.; Zhang, X.; Li, H.; Dulak, A.; Tomko, R.J., Jr.; Lazo, J.S.; Wang, Z.; Zhang, L.; Yu, J. Microrna-21 negatively regulates cdc25a and cell cycle progression in colon cancer cells. *Cancer Res.* **2009**, *69*, 8157–8165. [CrossRef] [PubMed]
144. Sarkar, S.; Dey, B.K.; Dutta, A. Mir-322/424 and -503 are induced during muscle differentiation and promote cell cycle quiescence and differentiation by down-regulation of cdc25a. *Mol. Biol. Cell* **2010**, *21*, 2138–2149. [CrossRef] [PubMed]
145. Kabir, F.M.L.; DeInnocentes, P.; Bird, R.C. Altered microrna expression profiles and regulation of ink4a/cdkn2a tumor suppressor genes in canine breast cancer models. *J. Cell. Biochem.* **2015**, in press.

Appendix 9

Accepted manuscript draft: Introduction to Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics

Martic-Kehl, M., Festing, M. F. W., Alvarez, C. E., Schubiger, P. A., Introduction to Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics; in Methods and Principles in Medicinal Chemistry series, Wiley-VCH, *In press*.

Wiley Series "Methods and Principles in Medicinal Chemistry",
Series Editors: R.Mannhold, H.Kubinyi, G.Folkers

Book Nr. 61 "Animal models in research and development of cancer therapy – Predictability of clinical efficacy"
Special Editors: M.I. Martic-Kehl, P.A. Schubiger

Introduction

Marianne Isabelle Martic-Kehl, Michael FW Festing, Carlos Alvarez, Jannis Wernery & Pius August Schubiger

Animal models in biomedical research

Modern biomedical research relies heavily on the use of laboratory animals, particularly mice, rats and fish which, in the UK in 2013, accounted for 94% percent of animals used in research. The research included fundamental studies aimed at understanding biological processes, the pre-clinical testing of potential new drugs and therapies, the development of diagnostic reagents and, in the case of monoclonal antibodies, the production of therapeutic agents themselves.

In all developed countries the use of animals should (and probably is) strictly regulated in order to minimise pain and distress. All research workers should be familiar with the "Three Rs", Replacement, Refinement and Reduction described in the book The Principles of Humane Experimental Technique (Russell and Burch, 1959). So where possible, non-sentient alternatives to the use of animals should be used as a "Replacement". If animals must be used, then "Refinements", such as anaesthesia and analgesia as well as enriched housing conditions should be used to minimise pain, distress or lasting harm, and the number of animals used should be "Reduced" to the minimum necessary to meet the objectives of the study.

There are continued, successful, efforts to develop alternatives to the use of animals. For example, large numbers of animals were once used for assaying many biological reagents such as hormones and vaccines. These have now largely been replaced by in-vitro methods such as direct immunological or chemical assays. Fundamental research uses large numbers of mice but also makes extensive use of cell cultures and tissues from animals, which have been humanely euthanased. An important "Refinement" has been the development of disease-free or so called "specific pathogen free" mice, rats, guinea-pigs, rabbits and cats and a few other species. These are free of clinical and sub-clinical infections which can cause problems if the animals are stressed by an experimental treatment. "Reduction" is achieved by good experimental design in which neither too many animals are used, which would be wasteful, nor too few which might mean that important reactions are missed.

In vivo research takes an important role in life science, particularly in preclinical drug development. The drug development process as it is standard nowadays consists of several sub-phases (research phase, pre-clinical development, clinical phases I-III), taking several years and give rise to costs between 50 Million and 2 Billion US Dollars (*DiMasi, Adams, Light*).

The drug developing process is highly prone to attrition. One critical step is the translation of pre-clinical animal research results to the clinic. In the last decade, it was frequently revealed in many research fields, that translation rates are minimal, non-existent or generally shrinking (*Sena 2007, Baker 2014, Arrowsmith*). A literature investigation by Thomson-Reuters has revealed that the success rates of development projects in phase II clinical trials have fallen from 28% to 18% between 2009 and 2010 (*Arrowsmith*). In more than half of the cases the reason for attrition identified was insufficient efficacy.

Animals in the drug development process – historic background

Animals as surrogate organisms for humans were already used in the 19th century for the purpose of understanding chemistry-based drug effects on physiological function. A first Pure Food and Drugs Act in the USA (1906) described official standards for drugs and proper labeling and prohibited the interstate commerce of unsafe drugs. In 1938 the Food, Drug, and Cosmetic Act further required proof of safety and authorized inspections as a consequence of the sulfanilamide-disaster in 1937 that killed over hundred people in the USA. The reason for the toxicity was that the syrup with the antibiotic sulfanilamide administered against coughing contained the toxic solvent diethylene glycol. The syrup was very popular with children due to the sweet taste of the diethylene glycol.

The first privately financed, nationally supervised and evaluated drug was Streptomycin, in 1945. However regulatory authorities and the movement towards supervised drug development was still restricted to the Anglo-Saxon world. At that time in continental Europe the view persisted that diseases were non-comparable processes, not suitable for statistical evaluation "since the treatment never concerns populations but only individual patients" (Virchow, 1847). In 1954, the German company Grünenthal patented the sedative thalidomide and launched it 1957 as Contergan® in West Germany. Since the company considered the drug particularly safe they marketed it as a sleeping pill for pregnant women, as well as against morning sickness in early pregnancy. However two years later first reports about nerve damage related to Contergan® appeared and 1961 Grünenthal had to withdraw the drug from the market when thousands of babies were born with extremity abnormalities. This tragedy known as the thalidomide-disaster, led to the first drug law in Germany (1961). The USA revised their existing drug law a year later to request proof of efficacy and sufficient pharmacological and toxicological results from animals to be granted a license for market authorization. This development founded the base for the drug development process, as it is known nowadays in developed countries. As up to now, it is not possible to fully replace live animals as human surrogates, however it is of high concern to perform such experiments in the most ethical way possible.

(http://www.fda.gov/AboutFDA/WhatWeDo/History/FOrgsHistory/CDER/CenterforDrugEvaluationandResearchBrochureandChronology/ucm114465.htm;
http://www.contergan.grunenthal.info/grt-ctg/GRT-CTG/Die_Fakten/Chronologie/152700079.jsp;
http://www.contergan.grunenthal.info/grt-ctg/GRT-CTG/Die_Fakten/Das_deutsche_Arzneimittelrecht_nach/152700071.jsp;jsessionid=B82E77391EBCF93D7DCD34660297CBF2.drp1; *Hildebrandt AG, 2004*).

In various official acts, for example of the Royal Society in the UK, the UK Department of Health or the US Department of Public Health, it is stated: "Virtually

every medical achievement of the last century has depended directly or indirectly on research with animals". Whether or not this statement has any proven evidence was unclear until 2008, when Robert Matthews has investigated on it. In his article, he came to the conclusion that even though the statement does not generally hold true, "animal models can and have provided many crucial insights that have led to major advances in medicine and surgery" (*Matthews RAJ, 2008*).

Indeed research using laboratory and domestic animals has underpinned many major advances in human medicine. Perhaps Louis Pasteur in the 19th century should be credited with the first use of scientific methods to develop new treatments for infectious disease. He used dogs and rabbits to develop methods of immunizing dogs and humans against rabies, a virus disease (although viruses were not known at that time), and sheep to immunize sheep against anthrax, a bacterial disease. His methods laid the ground work for the development of vaccines used now to control diseases such a polio, measles, mumps and rubella. Two infectious viral diseases, smallpox and rinderpest, a serious disease of cattle, have even been entirely eliminated from the wild and polio has nearly been eliminated. The first oncogenic retro virus was discovered in 1911 by Peyton Rous who found that cancer can be induced in chickens by injecting them with a cell free extract from a chicken tumor. Further studies of the biology of murine retro-viruses such as the Bittner mammary tumor virus and murine leukemia virus, meant that when HIV appeared, at least the biology of retroviruses was largely understood. Although infectious diseases have now largely been controlled in developed countries, new zoonotic diseases such as that caused by the Ebola virus which is maintained in wild animals in West Africa and various strains of the influenza virus present in wild and domestic birds remain a constant threat, in view of the rapidity with which diseases can be transmitted throughout the world. And antibiotic resistant bacterial remain a constant threat.

Transplantation of kidneys, hearts and other organs has saved many lives. This was made possible by the discovery of immunological tolerance by Peter Medawar in the 1950s. At that time it was known that skin grafts between two individuals would be rejected but it was assumed that this was a physiological problem. Medawar showed that reciprocal skin grafts between two different strains of mice are rejected. However, if lymphocytes of a donor strain were injected into baby mice of a recipient strain, treated adult mice of the recipient strain would then permanently tolerate grafts from the donor strain. This showed for the first time that graft rejection is an immunological rather than a physiological phenomenon, and that it can be controlled by immunological methods. The development of drugs such as cephalosporin to dampen the immune system, again using laboratory animals, has made organ transplantation possible.

The first chemotherapy was developed by Paul Ehrlich, who, in 1909 screened 606 chemicals for activity against the spirochete causing syphilis using rabbits infected with the organism, and found one, salvarsan, that was effective. For a time it was the most widely prescribed drug in the world. The development of new drugs now depends on an understanding of the biology of the disease, the identification of possible drug targets and the screening of large numbers of chemicals likely to interact with the target using, *in vitro* and *in vivo* methods involving research animals. Any potential new drugs will be tested in animals for safety and efficacy usually in mice, rats and dogs before proceeding to clinical trials.

The discovery of insulin in the early 1920s has saved many millions of lives. Banting and Best ligated the pancreatic duct of dogs and found that cells associated with the production of digestive enzymes degenerated, leaving islands of cells. These secreted the hormone later designated insulin, and they showed that it could be used to maintain diabetic dogs. The biochemist Collip developed methods of purifying it from porcine and bovine pancreases, using several thousand rabbits to assay it. Fortunately, although these insulins are different from human insulin, they are sufficiently similar to be effect in humans. Before then type I diabetes was usually fatal. For many years batches of porcine or bovine insulin had to be assayed using mice or rabbits. Frederick Sanger sequenced the insulin protein in 1955 and genetically modified human insulin is now produced in bacterial cultures and is assayed chemically.

Antibiotics have probably saved more lives than any other medical intervention. Penicillin was discovered by Alexander Fleming in 1928, but he was unable to isolate it and verify that it was effective. This was done by Ernst Chain and Howard Florey who were able to show that it was both effect and non-toxic in mice. They went on to develop a method of producing it on a large scale. Many other antibiotics have been discovered since then. For example, Selman Waksman discovered streptomycin in research involving mice, guinea-pigs and chickens.

Nutritional deficiency diseases are fortunately now rare in developed countries, though still a problem in some under-developed countries. Frederick Gowland Hopkins showed that young rats given diets of purified protein, carbohydrate, minerals and fat stopped growing, but when they were given a small amount of milk they grew. He postulated that there are substances required in the diet in minute amounts, later called vitamins. The vitamin that he discovered was designated vitamin A. His work coincided with that of Christiaan Eijkman who was attempting to find the cause of beriberi, a disease characterised by loss of feeling in the feet and difficulties in breathing. He injected the blood of soldiers, in hospital with beriberi, into chickens. But he noticed that some of the chickens which, were fed on scraps of the same diet of polished rice as the soldiers, were also getting sick. He showed that chickens which received unpolished rice remained healthy. The disease was caused by a deficiency of what we now call vitamin B1 (thiamine). Hopkins and Eijkman shared the 1929 Nobel Prize for their work.

The few examples given above demonstrate how, historically, animal research has contributed to the development of many areas of medicine. The development of monoclonal antibodies is a relatively recent advance which has resulted in a limitless supply of highly specific diagnostic reagents as well as many promising new therapeutic agents. B-cell multiple myelomas have been known in humans for many years, and it was known that they produced monoclonal antibodies, known as Bence-Jones proteins. In the late 1960s it was found, (Potter, 1972) that the BALB/c inbred strain of mice produced myelomas when injected i.p. with mineral oil. These myelomas were immortalised and could be maintained as permanent cell cultures. In 1975 Kohler and Millstein fused these myeloma cells with spleen cells from mice which had been immunised to sheep red blood cells and found that the “hybridomas” secreted antibodies to sheep red blood cells. They were able to select out individual hybridoma cells, each of which produced a monoclonal antibody. Subsequently, the immunoglobulin genes of the mice were replaced by the equivalent human genes

using genetic engineering so that human rather than murine monoclonal antibodies are produced. This avoids any possible problems associated adverse reactions to mouse proteins. Monoclonal antibodies (Mabs) are now used to treat several diseases such as some forms of cancer, and as anti-inflammatory treatments of diseases like rheumatoid arthritis and Crohn's disease. Many more are being tested in clinical trials. Due to their high specificity they are also widely used in the diagnosis of disease.

Other examples where animals have made important contributions include blood transfusion, joint replacements, reproduction and in-vitro fertilisation (allowing many otherwise infertile couples to have children), heart valve replacement, cancer and stroke. Moreover, veterinary medicine and human medicine are converging. Dogs and humans get many similar diseases such as cancer, obesity and type II diabetes. Dogs also get a number of hereditary diseases, in some cases as a result of many generations of selective inbreeding for breed characteristics which are inappropriate or incompatible with health.

Problems with translation of animal data to the clinic

Despite the impressive examples described above, many articles in scientific and non-scientific journals have criticized the quality and reporting of animal research in drug development in the last decades. For some diseases, animal models were found to have no predictive value for clinical applications. A mouse model developed to investigate cystic fibrosis, for example, turned out to show symptoms different from human patients, even though the same genetic modification was introduced (*Ameen, 2000*). Another example is the search for Human Immunodeficiency Virus (HIV) vaccinations using non-human primates as surrogate organism. Chimpanzees and macaques infected with the Simian Immunodeficiency Virus (SIV), a similar virus to the HI virus and of which the HI virus is assumed to have developed from, turned out to be responsive to various vaccination candidates, whereas none has been translated successfully to human patients so far (*Langley, Buckland, Bailey 2008, Bailey 2011*). Certainly, men are not 70kg mice and it is probably utopic to assume that the efficacy of any drug candidate can be 100% reliably predicted by the investigation of an animal surrogate, nevertheless, literature analyses in various fields of research have revealed a variety of potential causes apart from pure genetics for low predictive value of animal research data. Poor experimental planning, inappropriate statistical analysis and insufficient reporting are keywords frequently summarized in literature (*Pound, Bailey 2005, Hackam 2006, Hackam 2007, Knight, Matthews, Langley, Muhlhausler, PLOS Medicine editors, Eisen, Baker, Bertotti*) and it is well conceivable that the predictive value of animal research can be increased substantially by eliminating such methodological shortcomings.

A disease field where a lot of work has been done to tackle down potential reasons for translation failures is acute stroke. Literature analysis revealed that almost five hundred intervention candidates have shown satisfactory efficacy in animal models, whereas only three interventions are proven to be effective in patients suffering from acute stroke (*O'Collins 2006, Sena 2007*). For various interventions with positive outcome in animal models, meta-analyses were performed to investigate potential reasons for this high failure rate. Judging from a checklist with ten quality criteria, researchers found a trend towards overestimation of effect size in studies suffering from low quality (*Sena 2007*).

By analyzing quality shortcomings of the investigated studies, the investigators identified two main groups, which can be summarized as, first, general, stroke-independent and second stroke-specific shortcomings. The latter included the use of animal models (mostly mice) that did not reflect the general health state of an average stroke patient. Patients are often elderly, suffering from additional health problems like hypertension or diabetes (*Sena 2007*), whereas mice are young and healthy apart from the artificially introduced lesion to trigger stroke symptoms. Furthermore, other researchers identified discrepancies in the administration schedule of a particular drug candidate. Treatment onset occurred much sooner in animals (median ten minutes) than in patients (median five hours) (*Perel 2007*).
The other, more general group of quality shortcomings concerns the frequent neglect of study design and performance concepts in animal research, which are standard for clinical trials. These include random allocation of animals to test and control groups, blinded performance and assessment of study outcome and sample size calculation prior to study performance in order to guarantee a certain study power (which should by convention minimally be 80 to 90%) (*Dirnagl 2006, Sena 2007, Perel 2007*).
Similar issues with study quality were observed in Amyotrophic Lateral Sclerosis (ALS) research. Even though well-established animal models for ALS exist, translational success to the clinic is absent to date (*Scott 2008*). Like in the case of acute stroke research, it was found that the general quality of study design of the investigated literature was similarly poor, with lack of randomization reporting, blinding and power discussions (*Benatar 2007*).

It is widely appreciated that study power and bias can determine the likelihood that any research finding is true. It is intuitively clear that small studies of weak effects will be predominated by false positive discovery. However, many investigators appear to ignore the impact on reproducibility of the number of other studies of the same question that one is investigating. Further, it seems likely that many biomedical researchers do not appropriately consider how the probability of true findings is affected by the ratio of true- vs. non-relationships in the total number of tests. The first protection from irreproducible results was the conception of the scientific method, which has evolved into the development of various study and statistical analysis designs. Each field of scholarship tends to establish a set of gold-standard approaches to balance true discovery and feasibility. An underlying requirement of this approach is the deep involvement of statisticians in the earliest steps of experimental design (e.g., what is the question?; what are the properties of the data?; etc.). Unfortunately, it is common for investigators to focus more on study logistics and technical aspects. Many gather full datasets before carefully considering how the data will be analyzed. A significant proportion of biomedical investigators lack the necessary statistical expertise and it is difficult to appreciate how frequently they receive appropriate support from others [9]. One indication that reproducibility is a problem is the growing trend of biomedical journals to ask manuscript reviewers to specifically address statistical analysis (and advise if it is necessary to implement a formal statistical review). Similarly, the National Institutes of Health (NIH) is developing such initiatives, focusing on training scientists in experimental design, improving evaluation of grant proposals for aspects of study plans that most influence reproducibility (e.g., analytical design, blinding and randomization), and access to published and unpublished primary data [10].

The above-mentioned aspects of study design are very crucial not only when it comes to performance, but also when it comes to reporting of studies. Additional examples of experimental aspects that need clear specification in publications are details about investigated animals, their handling and housing. Information about the sex of investigated animals is often missing and generally, there is a bias towards the use of male animals versus females, because it is (unjustifiably) assumed that male animals are less prone to confounding hormonal cycle deviations (*Couzin-Frankel 2014*). It was even mentioned that the above mentioned sex bias (or better Y-chromosome bias) should be considered for cell studies equally (*Couzin-Frankel 2014*).

Insufficient reporting can bias the judgment of decision-makers regarding the translation of drug candidates from the bench to the bedside. Meta-analyses and systematic reviews provide powerful tools in such processes, as they estimate effect sizes of drug candidates by colliding all available published information. It has to be born in mind, though, that only published data can be analyzed in such studies. If researchers do not publish their findings or if they report study details sufficiently, such reviews will be strongly biased.
It has been proven, that the neglect of randomization and blinding leads to overestimation of effect size (*Bebarta 2003, Sena 2007*). It is difficult to judge, though, whether researchers not reporting on randomization in fact did not perform randomization or whether they just failed to report on it. On the other hand, reporting of randomization does not necessarily mean that randomization has been performed in a suitable way. It was revealed that many animal researchers do not seem to be aware of how (fully) randomized allocation of study subjects is performed appropriately (*Couzin-Frankel 2013*).

Generally, it was observed that studies with positive effect sizes (and often small sample sizes) tend to be published more frequently than studies with neutral or negative outcome. This phenomenon is referred to as publication bias in literature. Publication bias leads to distorted results of systematic reviews and meta-analyses. In the case of acute stroke research, it was estimated that approximately 30% of the pooled effect size for a certain drug estimated via a meta-analytic approach, was due to publication bias (*Sena 2010*). This corresponds to a substantial bias and it can be assumed that it this bias might reach similar dimensions in other fields of research.

Assuming the best technical and statistical practices are applied and reported, there are other aspects that can affect what we think we know to be true. Such factors can be psychological (e.g., self-deception [13] or effects related to professional incentives and personal motivations [3]) or philosophical. For instance, Karl Popper proposed a philosophy of science that addressed his notion that one cannot be certain that something is true – that we can only be sure that something is false [14]. The classical rule of inference is modus ponens (method of affirming): “If one knows the following two statements “If p, then q.” and “p.” are true, then one knows that “Therefore, q.” is true. But this approach requires absolute conditionals and generalizations in order to be true; such conditions can be challenging in real life. Put another way, the argument can be universally valid in logic but – if at least one of the premises is false in a given case – unsound. The alternative inference rule, central to Popper’s philosophy of science, is modus tollens (method of denying): “If p, then q. Not q. Therefore, not p.” This is now the dominant scientific approach. Popper’s method starts with a theory from which deductive reasoning is used to make conclusions and predictions.

Ultimately, critical testing of the predictions cannot prove that the theory is true; rather, testing can only falsify a theory. This is the reason why hypothesis-testing science is so dominant.

Animal studies in anti-cancer drug development

Failure rates of drug effects in the clinical test phase after successful animal experiments were reported to be highest in the field of oncology (*Hutchinson 2011*). In 2011, the licensing success rate for anti-cancer drugs reached 5%, in contrast to 20% for cardiovascular diseases (*Hutchinson 2011*).

Exemplarily, the situation of animal research quality within anti-cancer drug development shall be illustrated below. A systematic review about this drug category in preclinical animal experiments has revealed that less than half of all identified publications (n=232) were reporting on randomization (41%), blinded assessment of outcome (2%), allocation concealment (0%) and sample-size calculation (0.5%) (*Martic-Kehl*). Even though many articles were published in the last decade that emphasized on the importance of such study design features, there was no increase in reporting of them observable from the late 1990ies until 2011 (*Martic-Kehl*). The only exception was the increased incorporation of conflict of interest statements in more recent articles than in older ones (*Martic-Kehl*). This phenomenon can easily be explained by more extended author's guidelines of numerous scientific journals, which clearly request a conflict of interest statement.

It seems that external enforcement, e.g. by journal editors, are necessary to achieve an improvement of reporting quality and - presumably performance quality - of animal studies.

Furthermore, there was a tendency towards rather modest to inexistent drug effect for studies of higher quality compared to low quality studies (*Martic-Kehl*).

Anti-angiogenic cancer drugs represent a striking example of how animal models can mislead decision-making if they are not applied in a clinically relevant way. On the other hand, the anti-angiogenic cancer drug research proves well how clinically relevant animal models can provide important information not only about efficacy, but also about potential harmful side effects of drug candidates.

After the permission on the market, some evidence came up that certain anit-angiogenic drugs might trigger metastatic evasion of cancer cells in patients (*Ebos 2010*). Retrospectively, it was found that this phenomenon could have been foreseen by investigating metastatic cancer models (highly clinically relevant) in mice. Primary cancer models did not show corresponding results (*Ebos 2010*).

Similarly, the tumor location within the animal model can play a crucial role for predictive value as well. The easiest way of tumor inoculation into an animal and in addition the cheapest one is a sub-cutaneous injection at a prominent body part like shoulder or flank. An advantage of such locations is the easy follow-up of tumor growth. On the other hand, tumor cells completely grow out of their naturally occurring stromal conditions, which might influence their growth and reaction on cancer drugs crucially. For pre-clinical drug testing, it would therefore make sense to use orthotopic tumor models where tumors are inoculated at the organ of tumor cell origin. Eventually the animal species has to be considered carefully; a spontaneous dog or cat tumor may be genetically and behaviorally closer to human tumors than an induced human tumor in mice, which would not occur naturally.

Aim of the book

This book is dealing with animal models for cancer drug development, particularly for pre-clinical efficacy studies. It is thought to be complementary to the book "In vivo models for drug development" edited by Wiley in 2013 and with its focus on oncology drug development. This book aims at first to discuss all the relevant parameters, which are of importance to generate reliable animal results and second to show explicit examples of additional ways of preclinical tumor drug testing as well as to give practical recommendations to ensure maximal predictability of therapy efficacy on humans.

In the first chapter, ethical aspects of animal experimentation are depicted, as a base for the use of animal models in general.

The following three chapters contain general concepts, which are not specific for cancer research, but under all circumstances should be considered and implemented in cancer research as well. Those chapters broach the issues of study design and proper reporting, reporting bias and animal housing and handling.

In the fifth chapter clinically relevant animal models for anti-cancer drug development are discussed in detail.

The first subsection is discussing mouse models of advanced spontaneous metastasis and the discrepancies in efficacy between primary and metastasis models. This topic is followed by important lessons to be learned from human and canine trials to improve animal models in research through back translation.

Animal models for cancer are then expanded from rodents to species like cats and dogs and their potential use as spontaneous cancer models in research. The use of dog models is deepened in detail, depicting their potential advantage and complementary potential compared to models of induced tumors in rodents.

References

DiMasi, J. A. et al. (2003): *The price of innovation: new estimates of drug development costs*. In: *J Health Econ*. Bd. 22, S. 151-185

Adams, C. P. & Brantner, V. V. (2006): *Estimating the cost of new drug development: is it really 802 million dollars?*. In: *Health Aff (Millwood).* Bd. 25(2), S. 420-428

Light, D. W. & Warburton, R. (2011): *Demythologizing the high costs of pharmaceutical research*. In: *BioSocieties (Millwood).* Bd. 5, S.1-17

Arrowsmith J. Nature Reviews. 2011;10:1

http://www.fda.gov/AboutFDA/WhatWeDo/History/FOrgsHistory/CDER/CenterforDrugEvaluationandResearchBrochureandChronology/ucm114465.htm

http://www.contergan.grunenthal.info/grt-ctg/GRT-CTG/Die_Fakten/Chronologie/152700079.jsp

http://www.contergan.grunenthal.info/grt-ctg/GRT-CTG/Die_Fakten/Das_deutsche_Arzneimittelrecht_nach/152700071.jsp;jsessionid=B82E77391EBCF93D7DCD34660297CBF2.drp1
Hildebrandt AG. Pharmacology, Drug Efficacy, and the individual. Drug Metabolism Reviews. 2004;36:845-852
Matthews RAJ. J R Soc Med 2008 (1001) 95-98
Ameen N et al. Histochem Cell Biol. 2000;114:69-75
Langley G. RSDA 2009;1:161-168
Buckalnd G. In: Replacing Primates in Medical Research: An expert report by the Dr Hadwen
Trust, FRAME and Advocates for Animals. Available from science@drhadwentrust.org
Bailey J. ATLA. 2008;36:381-428
Bailey J. Alter Lab Anim. 2011;39:527-540
Pound P et al. BMJ 2004;328:514-517
Bailey J. Biogenic Amines 2005;19:235-255
Hackam DG. JAMA 2006;296:1731
Hackam DG. BMJ 2007;334:163-164
Knight A. AA TEX 2007;14:125-130
Muhlhausler BS et al. PLOS Biology. 2013;11:e1001481
The PLOS Medicine Editors. PLOS Medicine. 2013;10:e1001493
Eisen JA et al. PLOS Biology. 2014;12:e1001757
Baker D et al. PLOS Biology. 2014;12:e1001756
Bertotti A et al. JNCI. 2013;105:1426-1427
Sena ES. et al. TRENDS in Neuroscience 2007;9:433-439
O'Collins, V.E. et al. Ann. Neurol. 2006;59:467–477
Perel P. et al. BMJ. 2007;334:197-202
Dirnagl U. Journal of Cerebral Blood Flow & Metabolism. 2006;26:1465-1478
Benatar M. Neurobiology of Disease. 2007;26:1-13
Bebarta V. et al. ACAD Emerg Med. 2003;10:684-687
Couzin-Frankel J. Science. 2013;342:922
Sena ES. et al. Plos Biology. 2010;8:e1 000344
Couzin-Frankel J. Science. 2014;344:679
Hutchinson L., Kirk R. Nature Reviews Clinical Oncology. 2011;8:189-190
Martic-Kehl et al. unpublished results; Martic-Kehl, Buch Harald
Ebos JML, Lee CR, Cruz-Munoz W, Bjarnason GA, Christensen JG, Kerbel RS. Cancer Cell 2009;15:232-239
Scott S, Kranz JE, Cole J, Lincecum JM, Thompson K, Kelly N, Bostrom A, Theodoss J, Al-Nakhala BM, Vieira FG, Ramasubbu J, Heywood JA. Amyotrophic Lateral Sclerosis. 2008;9:4-15

Potter M. 1972. Immunoglobulin-producing tumors and myeloma proteins of mice. Physiological Reviews 52:631-719.

Russell WMS, Burch RL. 1959. The principles of humane experimental technique. Potters Bar, England: Special Edition, Universities Federation for Animal Welfare.

9. Ioannidis, J.P., et al., Increasing value and reducing waste in research design, conduct, and analysis. Lancet, 2014. 383(9912): p. 166-75.
10. Collins, F.S. and L.A. Tabak, Policy: NIH plans to enhance reproducibility.

Nature, 2014. 505(7485): p. 612-3.
11. Allison, D.B., et al., Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet, 2006. 7(1): p. 55-65.
12. Erickson, S., K. Kim, and D.B. Allison, Composite hypothesis testing: an approach built on intersecion union tests and Bayesian posterior probabilities. In Meta-analysis and Combining Information in Genetics and Genomics. 2010, NY: Chapman & Hall/CRC.
13. Weller, J.I., et al., A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. Genetics, 1998. 150(4): p. 1699-706.
14. Zaykin, D.V., S.S. Young, and P.H. Westfall, Using the false discovery rate approach in the genetic dissection of complex traits: a response to Weller et al. Genetics, 2000.154(4): p. 1917-8.

Appendix 10

Accepted manuscript draft: Dog models of naturally occurring cancer; in Animal Models for Human Cancer

Fenger J. M., Rowell, J. L., Zapata, I., London C. A., Kisseberth W. C.**,** Alvarez, C. E. (2015) Dog models of naturally occurring cancer; in Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics Methods and Principles; Medicinal Chemistry series, Wiley-VCH, *In press*.

Chapter Title: 8.4: Dog models of naturally occurring cancer

JM Fenger, JL Rowell, I Zapata, WC Kisseberth, CA London, CE Alvarez*

Within WILEY Series: "**Methods and Principles in Medicinal Chemistry**"
Editors: R. Mannhold, H.Kubinyi, G. Folkers

In: "**Animal Models for Human Cancer: Discovery and Development of Novel Therapeutics**"
Marianne I. Martic-Kehl, Pius A. Schubiger, Raimund Mannhold, Gerd Folkers, Helmut Buschmann
ISBN: 978-3-527-33997-6
248 pages

Joelle Fenger, DVM, Diplomate ACVIM (Oncology)
Department of Veterinary Clinical Sciences
The Ohio State University
Veterinary Medical Center
601 Vernon Tharp Street
Columbus, Ohio 43210
Ph: (614) 292-3551
fenger.3@osu.edu

Jennie Lynn Rowell, PhD, RN
Assistant Professor
Center of Excellence in Critical and Complex Care
College of Nursing
The Ohio State University
390 Newton Hall
1585 Neil Ave.
Columbus, OH 43210
Ph: (614) 292-4497
Rowell.14@osu.edu

Isain Zapata, PhD
Center for Molecular and Human Genetics
The Research Institute at Nationwide Children's Hospital
700 Children's Drive
Columbus, OH 43205
Ph: (614) 355-3600
Isain.Zapata@nationwidechildrens.org

William C. Kisseberth, DVM, PhD, Diplomate ACVIM (Oncology)
Associate Professor
Department of Veterinary Clinical Sciences
The Ohio State University
448 VMAB
1900 Coffey Rd.
Columbus, Ohio 43210
Ph: (614) 292-3551
Kisseberth.2@osu.edu

Cheryl A. London, DVM, PhD, Diplomate (Oncology)
Professor
The Thekla R. and Donald B. Shackelford Professorship in Canine Medicine
Department of Veterinary Biosciences
The Ohio State University
454 VMAB
1900 Coffey Rd.
Columbus, Ohio 43210
Ph: (614) 292-3551
London.20@osu.edu

***Communicating author:**

Carlos E. Alvarez, PhD
Principal Investigator, Associate Professor
Center for Molecular and Human Genetics
The Research Institute at Nationwide Children's Hospital and
Departments of Pediatrics and Veterinary Clinical Sciences
The Ohio State University Colleges of Medicine and Veterinary Medicine
700 Children's Drive, W491
Columbus, OH 43205
Ph: (614) 355-3600
Carlos.alvarez@nationwidechildrens.org

Alvarez.73@osu.edu

8.4

Dog models of naturally occurring cancer

Joelle Fenger, Jennie Rowell, Isain Zapata, William Kisseberth, Cheryl London and Carlos Alvarez

8.4.1

Introduction

For more than one decade, investigators have sounded the alarm that most biomedical research findings are false [1-4] and have asserted that this contributes to the very low success rate of translating preclinical research to new human therapies [5, 6]. While some of the problem can be attributed to aspects such as poor experimental design and analysis, it is also important to be aware of the limitations of specific animal models. When choosing animal models for drug target validation and development of therapies, it can be useful to begin by considering the weaknesses of the current paradigm. Two leading reasons for the frequent failure of clinical trials are i) insufficient understanding of the pathophysiology or the biochemistry of the targeted pathway, and ii) unacceptable toxicity or other side effects of therapeutic treatments. Our overarching goal here is to demonstrate how biomedical canine models can be used to dramatically advance our understanding of human biology, pathophysiology and clinical translation [7-10]. We stress that the canine model is not an alternative to the mouse. Rather, both models have unique strengths that complement each other and human translational investigation. We emphasize two concepts that can increase true discovery (and thus reproducibility): increasing power and reducing bias. Dog models achieve this by being a simplified and more homogenous version of human disease within breeds or breed groups, but having increased variation across breeds.

8.4.1.1

Animal models of human disease and the need for alternatives to the mouse

Animal models of human biology and disease can have vastly different utilities. Comparative biology and genomics allow investigators to determine the extent to which human physiology is conserved across phylogeny. Biological models are also necessary for the execution of

experiments that are not ethical in humans. The genetic simplification of a disease model and the control of as many experimental variables as possible can also be critical to biological understanding. The more complex a phenomenon is and the weaker the effect of individual contributions to a trait, the more important this is. Lastly, experimental flexibility, speed and cost are also prominent considerations when choosing the best organism in which to study a given question. In many cases, multiple organisms are used. One of the most common final aims is the creation of a rodent model that can be used for validation of human disease targets, mechanistic studies of pathophysiology and/or therapeutic studies. Presently, there are approximately a dozen models that have been established through characterization of their biology and the creation of genetic tools. Those experimental workhorses cover bacteria, protozoans, plants, yeast, nematode worms, flies, fish, frogs, birds and mammals [11].

Nobel Prize history provides a good measure of the impact of animal models on medical discovery [Official Nobel Prize web site, 12, see table of animal models at Foundation for Biomedical Research, 13]. The majority of investigations that resulted in Nobel Prizes in medicine (since the first in 1901) involved animal models. Since 1979 that holds for all awards except one that exclusively used a plant (1983, Barbara McClintock, for the discovery of mobile genetic elements). Most relevant to this review, there have been 29 award winners who used dog models (9 of those only used dogs). But if one compares the use of dogs or cats vs. mice or rats over time, a pattern is evident: from 1901-1973, the ratio was 35:24, respectively; from 1974-2013, the ratio was 7:53. Below we discuss how this gain in the use of rodent models for experimental investigation has recently begun to be offset by an explosion of veterinary and genetic research in dogs.

Although the tradeoff between experimental tractability and human relevance is clear [11], investigators have tended to believe that their options of animal models are restricted. Such notions were historically driven by study feasibility: mainly decreased heterogeneity of inbred lines, and reduced costs of small animals with short generation times. Until the last decade, the only mammal with a full spectrum of genetic tools was the mouse. Now it is possible to mediate targeted mutagenesis in diverse animals [14]. Similarly, in the past it was difficult to work outside a small number of animal models that had at least a known genome sequence and annotation. But now, since high throughput methodologies allow for sequencing on the scale of the human genome for approximately 1,000 USD, it is possible to rapidly initiate genomic

studies in any organism. Non-human primates offer unparalleled human relevance, but have great limitations for some experimental methodologies, have very high costs, and carry a heavy or unsustainable ethical burden. While bacteria, protozoa, yeast and invertebrates tend to be very flexible, rapid and inexpensive, they lack the relevance of vertebrate physiology [15]. In the area of cancer, there are many questions where mammalian models are necessary or highly preferable. Among the aspects that can be unique or very different in mammals are the existence of the mammary gland, reproduction, gestation, post-natal growth rate, warm-bloodedness and many other metabolic traits, biology of the skin and lungs, and particulars of the immune system [16].

Mice have the strengths of being evolutionarily close to humans and being relatively inexpensive to use [17]. Dozens of inbred strains have been created. Diverse genetic and molecular tools and reagents have been developed. This has led to the greatest body of investigation available for any animal model – encompassing genetics and genomics, physiology, genetic and other inducible disease models, mutagenesis screens and therapeutic drug screens. Rodents have thus become the predominant models across essentially all disease areas. However, scientists using mice as biomedical and pharmacological models of human disease must be aware of their potential weaknesses that can impact clinical translation [5, 6, 18]. A major limitation that receives little attention is the lack of bio-epidemiological understanding [19]. Murine cancer studies very rarely involve naturally occurring cancer. Instead, most mouse models involve the use of highly-penetrant oncogene transgenics. The normal roles of heterogeneity, complex genetic and gene-environment effects, aging, and early somatic mutation are thus completely removed from these models of tumorigenesis. A high proportion of murine cancer studies also lack other key aspects of human oncological relevance such as tissue-specificity, drug response, recurrence, metastasis and malignancy. Many studies, such as human xenograph models, require the use of immunoincompetent mice and thus ignore a central mechanism of cancer risk and development [20]. Consideration of these various limitations of rodent models highlights the complementary nature of veterinary animal models.

8.4.2
Advantages of spontaneous cancer models in dogs

By comparison to rodents, dogs are outbred, have more limited tools for genetic manipulation and are vastly more expensive to use (as they are larger and have longer generation times). That and trends in ethics have resulted in a relative decline in canine experimental models. Experimentation on dogs is widely believed to be unethical in countries where dogs are considered close companions (often viewed as family members) or indispensable working animals. However, dogs have many advantages that are becoming appreciated more and more broadly [9, 10]. The best known of those are powerfully-simplified genetics offered by hundreds of isolated breeds, 5- to 8-fold accelerated aging compared to humans, and health care levels second only to humans. Most major diseases of dogs are documented and many breed-specific risk factors are known (e.g., section 8.4.3 Dog cancer models; see Table 8.4.1& Figs. 8.4.1-2). In this section we highlight the unique traits that make dog models effective in the integration of biomedical genomics and translational medicine [10, 18, 21]. It should be noted that the various strengths are intimately associated and have synergistic effects on investigational power.

8.4.2.1

High level of evolutionary conservation with humans

Although humans and mice are more closely related to each other (divergence time ~74 MYA) than they are to dogs (~81 MYA), humans and dogs are more similar [22]. Dogs are genetically more similar to humans than mice are, and have about 650 Mb of genome sequence that was present in the common ancestor of dogs, mice and humans – but which has been lost in mice [23]. In addition, dogs are more similar to humans at the protein level [23, 24]. The reason behind these facts is relatively-accelerated divergence in rodents, a complex phenomenon that is generally associated with animal size and other traits [25]. There are nearly 400 inherited diseases characterized in dogs [26]. Essentially all of these resemble a similar condition in humans. At least ~73% of pure-bred dogs today have an inherited disease condition [27]. They represent all disease areas and most diseases seen in humans – including diseases of complex genetics such as diabetes and cancer. A caveat here is that the resources for conducting epidemiology of complex diseases in dogs are not in place. However, it is evident that the prevalence of diverse complex diseases ranges from similar to very different between humans and dogs. For instance, heart disease is much more common in humans. With regards to cancer, some have similar prevalence in both species (e.g., mammary), but others can be common in one

species and rare in the other [e.g., osteosarcoma (OS) is common in dogs and rare in humans; and prostate cancer is vice versa, 28; see Table 2]. In the cases where diseases have been dissected at the genetic level, the gene or pathway involved in dogs has been the same as in humans [9].

8.4.2.2

Reduced heterogeneity within breeds and increased variation across breeds

A great advantage of dog models is simplified genetics [9]. This stems from their evolution that involved i) a major population bottleneck in their domestication from wolves (about 11,000-16,000 years ago), and ii) a series of population bottlenecks upon the creation of breeds from thousands of years ago to the present [29]. The majority of the nearly 500 extant dog breeds were created in the last 150 years. Selection of variation was largely based on morphology and behavior, making dogs the phenotypically most-varied land mammals. It is estimated that only 5% of the variation in wolves was lost in the domestication of dogs [30]. However, the creation of each dog breed is estimated to have removed ~35% of the total variation in all dog breeds. Thus, dogs are genetically extremely similar within breeds, but are dramatically different across breeds. The extent of both of those traits is very exaggerated compared to the same in the main human ethnic groups – such as African, Asian and European – which vary ~0.1% within groups and ~10% across groups [31]. This relative genetic-simplicity provides several strengths. Each breed is on the order of 100-fold genetically simpler than the full dog (or human) population. The hundreds of breeds can be genetically clustered into 10 or so breed groups – with many members that have different levels of relatedness to each other. That makes dogs ideal for rapid mapping of traits to broad genome regions within breeds [as the extent of linkage disequilibrium, or LD, within breeds is on the order of 50-100 fold greater than in humans; see 9, 32], and fine mapping across breeds [33].

As a result of the evolutionary history of dogs under human domestication and their within-breed genetic simplification, dogs have vastly reduced heterogeneity in at least some complex traits. Thus the genetics and phenotypes are vastly more homogenous, and these two aspects synergistically increase the power of genome scanning of complex traits. The fact that dog breeds have extreme variation across diverse traits of all types further increases the strength of dog models for biomedical research. This can also be used to avoid biases common in studies

of single breeds, strains or species. As an example of heterogeneity in humans vs. dogs, let's consider the trait of height, which has a high level of heritability in both species (~80% for humans). A meta-analysis of 46 human genome wide association studies (GWAS's; discovery n=133,653; replication n=50,074) for height uncovered 135 associated markers that explained ~10% of height variation [32]. In contrast, multiple studies of hundreds of dogs from dozens of dog breeds revealed six genome loci (*IGF1, IGFR1, GHR, SMAD2, HMGA2* and *STC2*) that account for at least 53% of the size variation across all but giant breeds [34]. One study of 57 morphological traits in 915 dogs from 80 breeds showed that the majority of phenotypic variation for most of those traits could be explained by 1-3 loci [35]. A prominent question today is whether the genetic architecture and effect sizes of associated variants of canine diseases of complex genetics will be similar to that of canine morphological traits, but it is too early to generalize about this.

In the area of complex disease traits, a comparison between human and dog OS risk can be made. A large consortium of investigators conducted a multistage GWAS using 941 human OS cases and 3291 controls, and identified two genomewide significant loci (*GRM4*, Odds Risk, OR, of 1.57; and a gene desert on 2p25.2, OR 1.39) [36]. In contrast, Karlsson, Lindblad-Toh et al. conducted GWAS's of OS risk in three separate dog breeds (total n=543) and identified 33 genomewide significant loci that included prominent cancer genes *CDKN2A/B, AKT2* and *BCL2*. These were estimated to explain 57%, 53% and 85% of the phenotype variance in Greyhounds, Irish Wolfhounds and Rottweilers respectively [see section, 8.4.3.2.2 Osteosarcoma; 37]. The highest observed OR's for canine OS in those three breeds that all have a high prevalence of the disease were 1.36, 1.43 and 1.75. This could suggest the presence of additional loci in each breed that could not be mapped because the breed is fixed for the risk allele. This was found to be the case for seven of the loci, which were fixed in one or – for one locus – both other breeds. Other than that, there was no overlap between the OS loci in the three breeds, which included two related sighthounds. These observations represent an important caveat that has to be considered in the design and interpretation of dog studies. The reduced heterogeneity of breeds is associated with high levels of homozygosity that cannot be assayed in genetic mapping studies. Those regions may contain variation that contributes to essentially any trait. Additionally, the use of breeds with greatly reduced heterogeneity suggests that the findings will not necessarily be directly applicable across breeds. Two ways to mitigate this issue, and perhaps change it to a

strength, are i) comprehensive genotyping of all genetic variation (SNP and CNV; discussed below), and ii) analysis of multiple breeds, ideally including ones that are not closely related.

8.4.2.3

Potential for comprehensive genotyping

Besides the challenge of very high heterogeneity, another limitation to human complex genetics is the fact that mapping studies are not able to query a major portion of variation – mainly structural variation (SV) or DNA copy number variation (CNV). Structural variation is that which alters the structure or content of DNA, such as duplications, deletions, inversions and translocations; CNV refers to sequence that has variable copy number (arbitrarily defined as at least 50 bp), such as gain or loss structural variants [38]. SV/CNV are distinct from somatic genomic alterations because they are restricted to germline polymorphisms (with an arbitrarily established minor allele frequency commonly set at 1%). The extent of CNV as measured by number of variant base pairs in humans is 100 bases of CNV per single nucleotide variant (SNP). However, it is technically very difficult to genotype the majority of SV/CNV. Most GWAS's only capture a minute portion of SV/CNV, predominantly very large CNVs. Because much of SV/CNV involves repetitive sequences of diverse types, it is often difficult or impossible to develop specific genotyping assays. Whole genome sequencing in humans results in approximately 25% of the reads falling outside reference-templated assemblies. Those unmappable reads are almost entirely non-unique sequences that make up much of SV/CNV.

Because of the reduced genetic variation within breeds, dogs are ideal mammals in which to catalog all heritable variation in order to develop genotyping platforms that can detect most variants [39-41]. This could be done by identifying all SV/CNV and sequence variants in specific breeds, then creating an integrated map of those [18, 39, 40]. The use of novel long distance sequencing and analogous methods would allow determination of a high-quality phased assembly of breed-specific genomes. We estimate that the first version of this could be done with approximately 10-20 dogs of one breed. This would yield a set of tag SNPs that could detect most discerned haplotypes along with their known SV/CNVs. Such a SNP platform could be used for genome scanning or integrative genomics that incorporated an annotated catalog of all breed variants. The annotation of variants could include scoring of variants for likelihood of having functional or pathological effects. Annotation could also include classification of variants

into biochemical or genetic networks, or according to other types of biological relevance. Such annotation would yield a list of gene variants that may be implicated in diseases that are common to the breed under study. Implicated variants could be directly tested for association with that disease in that breed. Notably, this information would also be useful to consider the possible effects of candidate functional-variants that are present in genome regions that are fixed, or nearly fixed, and cannot be queried in genetic mapping studies [42].

8.4.2.4

Understanding both somatic and germline cancer genetics

Most of the strengths of dog models are relevant across all human disease areas, but there are aspects of cancer that differ from all other diseases [18]. Among those are the facts that cancer i) involves complex germline-risk genetics and somatic mutations, and ii) is also associated with broad types of environmental stimuli that include drugs and alcohol, diet, sunlight and diverse microbes. Another distinguishing characteristic of cancer is that it affects myriad cell types in essentially all tissues, and can involve almost any biochemical process of the cell. The recent explosion of molecular-genetic knowledge of cancer can be summed-up best as the realization that it is vastly more heterogeneous and complicated than we had ever imagined. This problem of tumor (including intra-tumor) heterogeneity is true within and across cancer types. Given those biological realities, canine breed models present powerful opportunities of simplified genetics and reduced disease heterogeneity. Compared to humans, it is vastly simpler to map traits and to conduct translational studies in dogs.

The understanding of human cancer genetics has exploded in the recent past. That exponential advance was made possible by the development of high throughput sequencing and its application to the sequencing of tens of thousands of tumor genomes, exomes or transcriptomes [supported by other ‘omics methods: DNA methylation, comparative genomic hybridization, chromatin IP, small and long non-coding RNA, proteomics, metabolomics, etc.; 43]. This has led to the identification of the majority of somatic driver genes in all common human cancers. Such understanding is very important to be able to go on to stratify cancers for biological dissection and therapeutic relevance. However, knowledge of germline cancer genetics in humans is almost a black box [7]. The reason for this is that the heterogeneity of human populations makes it impossible to map most genetic variation associated with complex

diseases such as cancer. Even for cancers that have been studied many times in the same general population – e.g., breast and prostate cancer in Western Europeans – only a small portion of the heritability of those cancers is explained. Furthermore, human complex genetics is at a very early stage and the vast majority of mapped associations are only defined as a general location – without knowledge of the causative variant (and any hint of disease mechanism) [44]. This complexity of common diseases is compounded by environmental risk factors. It is clear that our understanding of common diseases is embryonic. We must not only map gene variants, but will have to learn their biology, their interaction with other gene variants and their interaction with the environment (including with diverse microbes that are evolving rapidly).

Many of the successes in addressing the gap of in our understanding of human germline disease risk came from studying multigenerational families and isolated populations, and these continue to be important [45]. Such populations include the Finish, Icelanders, Utah Mormons, Amish, and Ashkenazi Jews. Examples of early contributions relate to the first breast cancer gene, *BRCA1*, which was mapped by family studies [46], identified in Utah Mormons [47] and found to reveal important insights in the Ashkenazi Jews [48] among other populations. However, neither this nor the new approach of sequencing-based discovery of "rare variants of strong effect" is suited for identifying most common variation associated with worldwide cancer risk. To date the only other alternative has been to increase the power of GWAS's in broader populations. It is no longer uncommon to have studies with 10,000 cases and as many or more controls. And these studies are still not revealing the majority of disease associated variants or heritability for complex diseases [49]. In terms of genetic simplicity, natural cancer in inbred rodents offers a way to address the gap in our understanding of germline cancer genetics. Such work is relatively rare in unmodified rodents because it would require great resources due to incomplete penetrance of risk variants and the requirements of advanced age. Interpretation of those studies is also complicated by the same reduced complexity that is desired: is the simplicity so great that the rodent strain is anomalous with regard to cancer risk? Rodent strains are selected for viability and fecundity in the presence of very little genetic variation – how are these lines different from rodent or human individuals and populations? In addition, the lifestyle and medical contexts of human cancer are very different from the laboratory setting of rodent cancer.

Pet and working dogs are attractive subjects for studying germline cancer genetics because they are naturally susceptible to most cancers present in humans, share environments

with their human companions, exist in a medical context and are genetically highly-simplified [see Table 8.4.3; 18, 24]. As with rodent lines, the knowledge gained from single breeds may not be generalizable across the population. This can be addressed by studying multiple breeds from different breed groups. Deep understanding of canine germline and somatic genetics associated with specific cancers and breeds – e.g., OS in Greyhounds and Rottweilers [37] – will lead to unprecedented knowledge of gene-gene and gene-environment interactions in a mammal. Another promising approach is to test for canine cancer-risk pathways in humans. Today this is complicated by lack of broad genotype data for humans, but it is likely that this will change in the next decade. Variant genes that are associated with canine cancer risk can be tested for direct association in humans, but they can also be used to implicate biochemical and genetic networks that can be assessed for human gene-gene/-environment interactions. Dogs could make major contributions in the latter area. There is very intense interest in this, but it is computationally impossible to test all potential combinations in large human datasets using conventional approaches at this time [50]. Canine models thus provide an opportunity to develop an understanding of germline contributions to diverse cancer traits. A complete picture will emerge when that is combined with somatic variation and environmental data collected in the veterinary setting.

8.4.2.5

Translational models

The same strengths of the canine model that set it apart from rodent and human studies – genetic simplicity in a natural population with highly advanced clinical care – also offer a unique solution to the gap between therapeutic research and development [7]. In dogs, it is possible to identify the germline and somatic cancer-drivers, and then to directly test targeted therapies on the same species. Such proof of concept studies could dramatically reduce the cost and time compared to similar studies in humans [10]. Below, we give examples of how dogs are specifically relevant to two human cancers and how the canine model is already being exploited for rapid development as a cancer model.

8.4.3

Dog cancer models

8.4.3.1

Canine cancer incidence

Comparative oncology integrates the study of naturally occurring cancers in animals with the study of human cancer biology and therapy [51]. Studying dogs with spontaneous cancer provides a valuable perspective distinct from that generated with other animal models because dogs naturally develop cancers that share similar characteristics to that of their human counterpart [8]; and this occurs in a population that, for the most part, is sharing many of the same environmental exposures with people. To this end, studies of dogs with cancer conducted over the past 30-40 years have significantly advanced the understanding and practice of human oncology in fields such as basic tumor biology [52, 53], tumor immunology [54, 55], radiation biology [56, 57], surgical oncology [58, 59], and systemic therapies [60] for a variety of cancers including OS, lymphoma, melanoma, and others.

Canine health foundations estimate that in the USA, ~8.5% of the canine population will be diagnosed with cancer this year (6 Million new cases/ 69.9 Million dogs) compared to ~0.5% of the USA human population [1,665,540 new cases in 2014/318,951,181 US population; 57, 58, 61, 62]. 27% of all dogs die from cancer, making it the leading cause of death [63]. Cancer is the leading cause of death in dogs over 10 years, with 50% of older dogs developing the disease [63-65]. Advances in the care of pet animals such as in better nutrition, vaccination for common infectious disease, leash laws that reduce automobile deaths, and the availability of more sophisticated veterinary care has allowed dogs to live longer. Those factors have increased the population of dogs at risk for the development of cancer and other age-related diseases. The rising prevalence of cancer in the pet animal population may be the result of an actual increase in cancer incidence. However, this may also reflect an increasing interest of owners to seek out advanced veterinary care or to participate in research studies [66].

Cancer in dogs and humans share many features including histologic appearance, tumor genetics, molecular pathway alterations, biological behavior, and similar responses to traditional treatment regimens such as surgery and chemotherapy (see Table 8.4.4). Similar environmental, nutritional, age, sex, and reproductive factors are associated with tumor development and progression in human and canine cancers. Pet dogs share the same living environment as their caregivers, potentially serving as epidemiologic or etiologic sentinels for the changing patterns of cancer development seen in humans [67-72]. Importantly, spontaneous cancers in pet dogs

recapitulate the biological complexity of human cancers in that they occur in the presence of an intact immune system and are characterized by tumor growth over long periods of time, inter-individual and intra-tumoral heterogeneity, development of recurrent or resistant disease, and metastasis to relevant distant sites [8].

The spectrum of cancers seen in humans is similarly diverse in dogs, including OS, NHL, melanoma, soft-tissue sarcoma, mammary, lung, head and neck, and bladder carcinomas [65, 73-78]. However, the incidence can vary dramatically between the two species [28]. While the most common human cancers are various epithelial types, most of those are relatively rare in dogs and the latter have higher rates of lymphomas and sarcomas. There is now a growing body of evidence from cross-species genomic analyses that demonstrate significant similarities between genomic profiles in canine and human cancers, providing support for the notion that these diseases are similar at a molecular [see section 8.4.2 Advantages of spontaneous cancer models in dogs; 52, 76, 79, 80-82]. For the more commonly studied cancers (e.g. OS and NHL), gene expression profiles and genomic alterations associated with cancers in dogs are highly analogous to those identified in their human counterparts [53, 73, 83-85]. Additionally, specific biochemical pathways known to be drivers in human cancers are also observed in various types of canine cancers, offering the opportunity to therapeutically target those mechanisms in dogs. That knowledge can then translate the findings more globally to human cancers driven by the same cancer-causing gene mutations. For example, approximately 30% of high grade canine mast cell tumors possess activating mutations in the tyrosine kinase receptor gene *KIT* (aka, *c-Kit*) which is associated with increased risk for the development of metastatic disease and shorter overall survival times [86-88]. While mast cell malignancies are uncommon in people, similar *KIT* mutations are found in human gastrointestinal stromal tumors (GIST) [89]. Clinical trials investigating the safety and efficacy of novel tyrosine kinase inhibitors targeting KIT in dogs with mast cell tumors have provided meaningful data correlating target inhibition with the mutational status of KIT and drug plasma concentrations, aiding in the translational development of this class of agents for KIT-driven malignancies in humans [90-92].

Partly due to the tremendous genetic advances (sequencing of the canine genome in 2005, development of methods to capture genetic variations on a high-throughput genomewide scale, etc.), the last 10 years have seen an explosion of canine cancer research. In fact, the publications related to canine cancer from 1970-2000, had an annual growth rate of 1.33% (total growth of

~40%). Compare that to 2001-2013 where the annual growth rate was 8.14% (total growth of ~98%; see Fig. 8.4.3). In the past, the predominant focus of canine cancer research was on the most common cancers in dogs, some of which were rare in humans (e.g., OS). Presumably because mammary cancer is rare in spayed dogs, comparative oncology of canine mammary cancer has been relatively limited. However, there has been a recent surge in study of canine models of breast cancer. As with humans, mammary tumors in dogs are the most common tumor type in females and are significantly associated with somatic mutations in *BRCA1* and *BRCA2* [93]. Additionally, recent research has recognized the many similarities regarding the morphology, biological behavior, and clinical course of mammary tumors in both species [94]. We thus expect the study of dog mammary cancer to grow rapidly. Below we focus on two types of cancer for which the dog arguably recapitulates the human disease better than any other animal due to natural disease, increased genetic variation compared to inbred rodents, and the relative richness of epidemiological and clinical data. We highlight the potential of the canine model to vastly accelerate our understanding of cancer genetics and treatments.

8.4.3.2

Genetics of breed-specific cancer models

8.4.3.2.1

Lymphoma

Lymphoma refers to a heterogeneous group of disorders involving monoclonal proliferation of malignant lymphocytes [95]. Human lymphoma is commonly divided into two categories: Hodgkin's Lymphoma (HL) and Non-Hodgkin's Lymphoma (NHL). Together, 5-year survival rates for lymphoma are very high, ranging from ~70-90% [95, 96]. HL is characterized by the presence of a Reed-Sternberg cell and is typically subdivided into classical HL and nodular lymphocyte-predominant HL [96]. HL is rarer than NHL but has a better 5-year survival rate. Alternatively, NHL includes more subtypes then HL and the literature has reported an increasing incidence within the past 30 years [95] that has appeared to stabilize more recently [97]. NHLs represent 5% of all new cancer cases and are the fifth leading cause of cancer death, and the second fastest growing cancer in terms of mortality. Incidence rates of NHL are particularly high in western societies and for most subtypes of NHL, higher in men than women [98-100]. In dogs, NHL accounts for approximately 10% of all malignant tumors [83% of all haematopoietic

malignancies; 53]. NHLs have been directly compared in humans and dogs many times, yet our understanding of the known canine incidence still comes from the 1970s [78]. Similar to humans, the proliferative activity of individual NHL can vary tremendously, with tumor mitotic activity and miRNA diversity serving as possible clinical indices of aggressiveness in dogs [101, 102]. Further, activating pathway processes in lymphoproliferative diseases are shared between the two species [101, 102]. There is evidence suggesting that the tumor microenvironment, activating pathway processes, clinical, cytological, and immunophenotypic properties are similar in the dog providing a robust model of the human disease [101-106].

The most common type of NHL is diffuse large B-cell lymphoma (DLBCL), representing over 44% of Lymphomas [107]. B-cells are lymphocytes that confer efficient and long-lasting adaptive immunity by the generation of high-affinity antibodies against microbial antigens [108]. These cells form an essential part of the humoral immune response and play a central role in overall immune logic [108]. The World Health Organization (WHO) recognizes 4 subtypes of DLBCL: 1) DLCBL not otherwise specified (NOS), 2) DLCBL with predominant extranodal location, 3) large cell lymphoma of terminally differentiated B-cells and 4) Borderline cases [109, 110]. However, despite these classifications, diagnosis is complex and often does not provide a clear homogenous subtype and topographical pattern [111]. The degree of variability within histologically identical DLBCLs prompted attempts to molecularly subcategorize the tumors [112]. Subsequent to the WHO categorization, a seminal work, using gene expression profiling (GEP), identified that DLBCL-NOS (the largest grouping of DLBCL) could be divided based on well-defined genetic signatures that have clinically predicative utility that has been replicated multiple times [112, 113]. Activated B-cell (ABC), germinal center B-cell (GCB), and primary mediastinal B-cell lymphoma (PMBL), are grouped based on the differentiation and maturation of distinct B-cells at separate stages [114]. Survival for 5-years based on subtype is roughly 30%, 59%, and 64%, respectively [112, 115].

By far the most common molecular subtype of DLBCL, GCB DLBCL typically occurs in children and young adults [116]. The GCB subtype is characterized by a much more favorable outcome and a spectrum of genetic aberrations, which include the t(14;18)(q32;q21) translocation, deletion of *PTEN*, amplification of the miR-17-92 cluster, or *TP53* mutations [117]. Expression patterns associated with the GCB subtype include markers of germinal center differentiation, such as *CD10* and *BCL6* [117]. While *BCL6* is highly expressed in the GCB

subtype, it is rarely expressed in ABC DLBCL. This transcription repressor has many critical roles that contribute to innate and adaptive immunity. Under normal physiological conditions, germinal centre B-cells share some characteristics of tumor cells including that they proliferate rapidly, evade growth checkpoint controls, and tolerate ongoing genomic instability [118-120]. During the humoral response, *BCL6* functions as a master regulator of the GCB phenotype. Germinal Centres emerge in the secondary lymphoid organs upon B-cell activation and provide the setting for massive clonal expansion and immunoglobulin somatic hypermutation (SH). Immunoglobin SH is an adaptive measure that allows mutations in the variable regions of immunoglobin genes to provide more antibody affinity maturation [See review by 118]. Although this is essential to permit immunoglobulin mutagenesis and maturation, germinal centre B-cells are prone to malignant transformation. B-cells that generate high-affinity antibodies are then selected to undergo terminal differentiation to memory cells or long-lived antibody-secreting plasma cells [see Figure 8.4.4; 118].

In contrast, the ABC subtype has significantly worse survival and is characterized by a distinct genetic background that includes the t(3;14)(q27;q32) translocation, trisomy 3, deletion of the *INK4A-ARF* locus, *BCL2* amplification and constitutive activation of the nuclear factor kappa B (NF-κB) pathway with high expression of NF-κB target genes [117, 121]. Interestingly, both GCB and ABC express *BCL2* (oncogene), although there is a four fold higher expression in most ABC DLBCLs compared with GCB DLBCLs [113, 121]. In addition, ABC lymphomas arise from cells that have completed SH and therefore contain static immunoglobulin heavy chain variable region sequences [122]. ABC DLBCL centers on MYD88, the key-signaling adapter in the Toll receptor pathway. Somatic mutations in the MYD88 TIR domain occur in 39% of ABC DLBCL tumors, and are the most frequent oncogenic lesion described in this subtype [123].

Canine DLBCL is an aggressive malignancy demonstrating significant overlaps with the human disease and is considered curable in less than 10% of dogs [124]. Just as in humans, DLBCL is the leading identified histotype (44.4%) and is of primary B-cell origin [82, 107]. One work explored the subtypes of DLBCL NOS in dogs, comparing their findings with human tumors. Similar to humans, differential expression split the groups into ABC-like from GCB-like [82]. Furthermore, these canine-specific "ABC/GCB" discriminating genes, while different from the human “ABC/GCB” gene list, were involved in the same pathways and processes (e.g., NF-

kB signaling and B-cell receptor signaling). And using the human gene list, canine and human lymphomas clustered according to subtype, not species [82]. Recently, *NF-KB* expression in DLBCL was compared using a principle components analysis and hierarchical clustering. Human and canine DLBCL grouped together and separately from the human and canine healthy tissue [53]. A clinical trial of dogs showed that using a peptide that effectively blocked constitutive *NF-κB* activity *in vitro* markedly reduced the mitotic index in 50% (3/6) of the ABC DLBCL canine participants [125].

Interestingly, *BCL6* (which is highly expressed in GBC DLBCL) is rarely expressed in either canine DLBCL subtype and this has been confirmed by others [82, 125, 126]. Collectively, these studies identify molecular similarities in human DLBCL that suggest pet dogs are a highly representative model of DLBCL for future studies, including therapeutic clinical trials [82]. It also provides evidence that dogs do have unique aspects of tumor biology that, as our understanding develops, will improve knowledge of cancer initiation, progression, and maintenance. In addition to the *BCL6* difference in expression, clinical case studies abound with variations in the traditional presentation of lymphomas in dogs, only highlighting the complexity of this disease and potential for dogs as a model [127-130]. Further, findings in canine DLBCL can provide new targets for the human disease. For instance, tissue factor pathway inhibitor 2 (*TFPI-2*) is a tumor suppressor involved in invasiveness inhibition that has been explored in many cancers, but not lymphomas [124]. Recently, a study that looked at epigenetic silencing in canine DLBCL found high-frequency epigenetic dysregulation of *TFPI-2* that correlates with its reduced mRNA expression [124]. This provides one clear target for human analysis.

8.4.3.2.2

Osteosarcoma

Dogs have served as an epidemiological and pathological model of OS for decades. However, it is only more recently that the full potential of the canine application to human OS has been recognized – particularly in its relevance to genetics and biomarkers [comprehensively reviewed in 73]. OS is a high-grade primary bone neoplasm of mesenchymal origin [131]. The parallels between canine and human OS are significant in their clinical presentation, biologic behavior, histology, conventional treatments, and shared biological targets for investigational treatments [132]. In humans, primary bone tumors are rare and account for only 0.2% of all malignant

tumors [133]. OS has a bimodal age distribution that peaks at 15 to 19 years and ~70 years [16] with approximately 60% of tumors occurring in patients under 20 years of age, and ~10% in patients older than 60 years [134]. For the majority of cancer diagnoses, 77% are made in persons ≥ 55-years of age [135]. However, OS is the third most frequent cause of cancer in adolescents and represents over 56% of all bone tumors. The highest incidence in the US occurs in African American and Hispanic/Latino adolescents, and the lowest incidences are in non-Hispanics of European descent, Asians and Pacific Islanders [136]. Encouragingly, although the incidence of OS has not changed appreciably since 1975, the overall mortality has decreased [136]. Children aged birth-19 years diagnosed with OS from 1975-1979 had a 5-year survival rate of 45% [136]. With experience and multiple clinical trials, a chemotherapeutic regimen was defined that resulted in an increase in overall survival to 71% of those diagnosed between 2003-2009 [136, 137]. Now, those who survive 5-years have a 90% chance of surviving to at least 15-years [136]. However, new regimens in humans have failed to further improve outcomes since 1987 [137]. Thus, the canine model of OS is poised to help develop new therapeutic targets and improve outcomes. Dogs have already had a significant impact on OS as it was in dogs that the first limb-sparing techniques were developed [10] The incidence in dogs has been estimated to be at least 13.9/100,000 [10], but also reported as 52/100,000 [138] compared to actual US incidence of 0.947/100,000 in humans in 2014 [3,020 Cases/318,673,000 USA Population=0.947 Cases Per 100,000 individuals; 139, 140, 141].

A key clinical difference between humans and dogs is the peak onset of the disease in humans is at a time when growth is rapid, whereas in dogs, the peak incidence is later in life after closure of the growth plates [132]; however, a bimodal overlapping incidence is shared by the two species [10, 142]. OS primarily affects the metaphyseal of the weight-bearing regions of the long bones of the appendicular skeleton (shoulder girdle, arms, legs, etc.) accounting for 72% and 94% of OS in dogs and humans with the small remainder of tumors arising in the axial skeleton [including the flat bones of the skull, ribs, vertebrae, sternum, and pelvis; 122, 131, 139, 143-145]. Interestingly, proximal humeral location is a significant negative prognostic factor in both canine and human appendicular OS [73, 146]. Other prognostic factors associated with survival of OS shared by dogs and humans include tumor grade, tumor mitotic index, presence of metastasis, the use of adjuvant and/or neoadjuvant chemotherapy, postoperative infection at limb-sparing surgical sites, age of onset, and high serum alkaline phosphatase (ALP) and lactase

dehydrogenase (LDH) [73, 142, 143, 146]. In both species, a slight predominance in males over females has been reported [73, 142].

The exact etiology of OS remains obscure primarily because of the difficulty in isolating a pure population of osteocytes from residence within a mineralized matrix [now thought to be the mechanosensory cell in bone which has a major role in the regulation of bone formation and resorption; 147]. Recent work suggests that osteocytes can serve as OS progenitors as identified by expression of dentin *matrix phosphoprotein 1* (*DMP1*; a marker of osteocytes) in multiple murine, human, and canine OS cell lines [148]. To assess the tumorigenic potential of osteocytes, an immortalized murine osteocyte cell line was injected into two sites on mice and both locations developed tumors consistent with OS [148]. The strength of this etiology work will continue to be explored using multiple models, as will more traditional genetic research. These various approaches will lead to the validation of targets for clinical trials and pharmacogenetics.

Although expression of canine and human OS had been previously considered separately, the first study to consider both human and canine OS expression together replicated former classifications of short/long-term survival in another set of dogs and in 5 independent human datasets [52, 149]. Because of these similarities, dogs make an outstanding pre-human trial clinical model for the targeting of pharmacogenetic therapies [150]. For example, a potential clinical target is *Survivin,* a gene whose expression is elevated in human and canine cancer [151]. This gene is a member of the inhibitor of apoptosis (IAP) gene family, which encodes negative regulatory proteins that prevent apoptotic cell death [152]. *Survivin* expression has been reported to be an indicator of poor prognosis, low apoptotic index, poor differentiation, high proliferation activity, and enhanced resistance to cisplatin-mediated chemotherapy [153]. *Survivi*n attenuation in canine OS cells demonstrated inhibition of cell-cycle progression, increased apoptosis, mitotic arrest, and chemosensitivity to significantly improve *in vivo* tumor control [154].

Another clinically relevant target is the *Notch* signaling cascade. This pathway is crucial for the development of multiple organ systems including bone development and both osteoblasts and osteoclasts [155]. In human OS, the *Notch* target gene *HES1* is associated with cell proliferation, invasion and metastasis, and has been suggested to be a prognostic biomarker [156]. Some of these findings were replicated in dogs, but interestingly, there appears to be reduced *HES1* expression despite elevated expression of other *Notch* signaling targets in the most aggressive OS tumors [157]. These findings indicate that *HES1* expression, while a good

prognostic indicator, may not be an appropriate surrogate marker of *Notch* signaling [157]. Such findings are invaluable to future translational work and might have been missed without use of the canine model of OS. Thus far, shared aberrant gene expression in both canine and human OS includes *p53, RB1, PTEN, HER-2, MET, STAT-3, mTOR, ezrin, PDGFs/PDGFRs, MMP2/9, miR-134, miR-544* [158-168, for a list of developmental pathway abberations in OS see 169].

OS in companion animals is generally considered a disease of large and giant breed dogs [63, 170]. A hallmark study of canine OS in 1966 reported an OR of 185 for giant breeds (> 36 kg), and an OR of 13 for medium sized dogs (18-36kg) when compared to dogs weighing <10kg [171]. A follow-up study in 1998, conducted on 1462 cases of canine OS, found that ~30% of all OS cases were attributed to dogs weighing > 40 kg and only 5% of their tumors occurred in the axial skeleton [17]. Coupled with this, only 5% of OSs occurred in dogs weighing <15 kg, but ~60% of their tumors originated in the axial skeleton [17]. This suggests that determinants of OS may differ for large and small breeds, although increasing height and weight are the most predictive factors for OS in dogs [172]. Interestingly, height and weight have been found to be significant risk factors in human OS as well [145, 173]. In fact, compared to average birth-weight subjects (2,665-4,045 g), individuals with high birth-weight (≥ 4,046 g) had an increased risk of OS (OR 1.35) as did individuals >51st height percentile (OR 1.35) [174].

Insulin-like growth factor (IGF1) and *Insulin-like growth factor 1 receptor (IGF1R)* have been associated with height and weight. It was shown that *IGF1* null mice are 40% smaller than littermates, while *IGF1R* null mice are approximately 55% smaller and die at birth [175]. In dogs, a non-synonymous SNP (chr3:44,706,389) in *IGF1R* changes a highly conserved arginine at amino acid 204 to histidine and is predicted to prevent formation of several hydrogen bonds within the cysteine-rich domain of the receptor's ligand-binding extracellular subunit [176]. This SNP is associated with breed average height at withers of ≤ 10 inches and ~10 lb weight [34, 176] and deletion or mutation in humans causes severe short stature or non-syndromic children that are small for their gestational age [177-179]. Further, a single *IGF1* SNP haplotype is common to all small breeds and nearly absent from giant breeds, with the size of different breeds correlating with reduced *IGF1* plasma levels [180, 181].

IGF1R expression in human OS has been associated with tumor metastasis and poor prognosis, and represents an appealing therapeutic target [182]. Investigation of *IGF1R* expression in canine OS tissues and cell lines showed that *IGF1R* was expressed in 71% of the

samples and that dogs with higher levels of *IGF1R* expression (47% of cases) had significantly-decreased survival when compared to dogs with lower *IGF1R* expression [183]. The expression of the *IGF1R* gene is negatively regulated by a number of transcription factors, including the *WT1* and *p53* tumor suppressors that are known to be aberrantly expressed in both canine and human OS [184]. *IGF1* has also been associated with a genetic susceptibility for OS [185]. The role *IGF1 and IGF1R* play in growth, development, and OS is clearly important but still unclear [reviewed in 180]. This evidence suggests a crucial role for genes that govern growth in OS development. The dog provides a powerful model from which to uncover these relationships; for example, the average weights of Chihuahuas and English Mastiffs differ by 65-fold.

As discussed above (8.4.2.2), the first GWAS's for canine OS susceptibility were recently conducted on the Greyhound, Rottweiler, and Irish Wolfhound [42]. A total of 33 loci were identified and top candidates in each breed included prominent cancer genes such as *CDKN2A/B, AKT2* and *BCL2*. The top Greyhound candidate locus was fine-mapped to 15kb between *CDKN2B-AS1/ANRIL* and *CDKN2A/B.* An enhancer screen in the human OS U2OS cell line then narrowed the causative variant to a highly-conserved single nucleotide predicted to lie within a PAX5 binding site. The latter transcription factor offers an excellent candidate mechanism as it is a known regulator of bone biology. Fortuitously, the causative variant is the peak SNP marker in the Greyhound GWAS, and thus much evidence already exists for its allele frequency in diverse breeds. It is fixed in Rottweiler (0.97 frequency) and Irish Wolfhound (0.95), but present in lower frequencies in a panel of 28 different breeds. In a group of eight large (a major OS risk factor) and high OS-risk breeds (Rottweiler, Irish Wolfhound, Leonberger, Pyrenees, Mastiff, Labrador Retriever, Great Dane, and Golden Retriever), the risk allele is slightly more common in cases. In the Leonberger and Great Pyrenees it is weakly correlated with OS, but no association was found in the Mastiff, Labrador Retriever, Great Dane or Golden Retriever. Curiously, the OS risk variant is present in approximately 50% of several dozen breeds studied to date, but has not yet been detected in the more limited numbers of wolves that have been genotyped to date. As the *CDKN2A/B* locus is a pan-cancer tumor suppressor, it will be interesting to learn whether the dog risk variant affects OS or broad cancer risks in mice. This study hints at the power of combining germline and somatic genetics to identify pathophysiology of dog cancer. This type of deep biological understanding will lead to the development of targeted therapies and their translation in the same species.

8.4.4

Preclinical and veterinary translational investigations in dogs with cancer

8.4.4.1

Preclinical investigations in dogs with spontaneous cancer

The development and approval of new cancer drugs is a lengthy, costly, and frequently unidirectional process [186, 187]. Novel agents are assessed in conventional preclinical models of efficacy and toxicity prior to moving into human clinical trials; however, most new cancer drugs that enter human clinical trials fail to reach approval, largely because of unanticipated toxicity or lack of efficacy that was not predicted in rodent models of cancer [188, 189]. Based on our knowledge of the complexity of cancer, it is not surprising that many models fall short of being predictive. Translational studies investigating new drugs, devices and imaging techniques in pet dogs with cancer can overcome many of these shortcomings, and assist in the transition between conventional preclinical models and human clinical trials.

Mouse models of cancer have proven to be excellent tools for dissecting the biology of molecular pathways involved in cancer development and progression; however, they frequently do not truly recapitulate the biological features that define cancer in humans, including genomic instability and the heterogeneity of tumor cells within a complex microenvironment [190]. Furthermore, conventional mouse models fail to recapitulate the complex biology of cancer recurrence and metastasis integral to outcomes in human patients. Naturally occurring cancers in dogs have an intrinsic advantage as a model for human disease in that they mimic and represent biologically complex conditions in a way that is not possible using other animal models. In many cases, cancers in dogs are described in the same language as their human counterpart and can be classified according to histologic and/or clinical staging systems analogous to that used in human cancers [e.g. National Cancer Institute Working Formulation, World Health Organization histopathological classification and clinical staging system for domestic animals with lymphoma; 191, 192]. Given their large size, the evaluation of novel therapeutic approaches (drugs or devices) in pet dogs can answer important questions regarding relevant drug exposure that are often inadequately considered in mouse models. Additionally, the similarities between humans and dogs with respect to their size, tumor biology and anatomy provide an opportunity to

engineer devices for limb sparing or prosthesis and optimize surgical interventions that are challenging to recreate in other animal model systems [51, 74, 193].

The ability to rapidly advance therapeutics for rare human malignancies such as pediatric OS is limited by the low incidence of these diseases in humans [194]. In contrast, OS is at least ten times more prevalent in dogs, providing a significantly larger patient population in which to evaluate new treatment strategies. Importantly, pet dogs represent a large population size and their owners are highly motivated to seek out new treatment options for their pets, which provide a unique opportunity to sufficiently power clinical trials, including the assessment of new drugs [8, 51]. Serial tumor biopsies and repeated collection of body fluids (serum, whole blood, urine) from dogs before, during, and after exposure to an investigational agent allows for evaluation of clinical and biological endpoints (e.g., pharmacokinetics and pharmacodynamics) that can be linked to drug exposure, surrogate imaging or circulating biomarkers, and therapeutic response in ways that are often difficult or unacceptable in human trials [195, 196]. Pet owners are often willing to permit autopsy, which is crucial not only for assessment of tumor control, but also treatment-related toxicity. Lastly, the naturally shorter lifespan of dogs compared to humans, coupled with short survival times achieved with current treatments for canine cancers accelerates the pace at which clinical trials in dogs can be conducted and allows for more rigorous evaluation before translation into new human trials [195]. The compressed course of cancer progression seen in dogs allows timely assessment of novel cancer therapies and permits less costly outcome determinations, such as time to metastasis, local recurrence and survival.

Because no established gold standards exist for the management of cancer in dogs, the evaluation of novel therapeutics is possible in less advanced or less heavily pretreated pet populations, compared to human cancer patients participating in early-phase human trials. Unlike human clinical trials, investigational studies in dogs are not constrained by traditional Phase I, Phase II, and Phase III trial designs allowing new forms of treatment (especially investigational single-agent trials) to be offered to pet dogs prior to conventional therapies or during the period of minimal residual disease. This provides the unique opportunity to evaluate single-agent activity or combination therapies earlier in the drug development process.

Organized cooperative efforts are now in place in the United States to facilitate the inclusion of pet dogs with naturally occurring cancer into the development path of new cancer drugs. The Veterinary Cancer Society and the Veterinary Cooperative Oncology Group have led

efforts to encourage multicenter collaborative veterinary oncology studies and to enhance case accrual and facilitate clinical trials. Additionally, the Comparative Oncology Program (COP) of the National Cancer Institute (NCI) at the National Institutes of Health has established the Comparative Oncology Trials Consortium (COTC) to conduct rigorously controlled and focused preclinical trials of new cancer drugs intended to inform the design of human studies [8, 51, 197]. The COTC functions to facilitate the design and execution of clinical trials in dogs in collaboration with extramural academic comparative oncology centers, the pharmaceutical industry, and nongovernmental groups interested in cancer drug development. A multi-institutional Pharmacodynamics Core was established through the COTC to provide an infrastructure within the veterinary research community to support the development, validation, and assessment pharmacokinetic and pharmacodynamics end-points within COTC trials [198]. Finally, the Canine Comparative Oncology and Genomics Consortium (CCOGC) has established a national canine cancer biospecimen repository as a resource to facilitate comparative genomics and the identification of valid tumor targets in canine cancers to aid in preclinical drug development [51]. This national infrastructure is now able to i) support and facilitate the implementation and conduct of multi-institutional studies, ii) more directly engage the veterinary oncology community, and iii) respond to the needs of the pharmaceutical community to better inform the drug development path of new cancer drugs.

8.4.4.2

Conduct of preclinical and translational studies in pet dogs with cancer

The value of including pet dogs with cancer into preclinical studies intended to support the development of human cancer treatment strategies is significant. Studies utilizing dogs with cancer can inform many different aspects of the human preclinical drug development process (see Fig. 8.4.5). Importantly, these studies are facilitated by the ability to use the “species in kind” approach. That is, these studies can be conducted with the knowledge of the drug toxicities and pharmacokinetics derived from studies performed in healthy animals (laboratory dogs) of the same species. Preclinical studies in tumor-bearing dogs can be used to address questions of toxicity and pharmacokinetics in patients with cancer, treatment schedule, pharmacodynamics assessment and endpoints, efficacy, and others [91, 92, 197].

Clinical trials performed in companion animals with the goal of informing human oncology are considered to be preclinical studies with respect to human drug development [199]. The term clinical trial has a broader definition in veterinary medicine and is used to describe any clinical research study that enrolls client-owned companion animals. Recent efforts by key opinion leaders from the pharmaceutical and biotechnology community, academia, and regulatory and federal agencies have proposed guidelines for the conduct and oversight of preclinical translational studies that include pet dogs with cancer intended to support the development of human cancer drugs or treatment delivery devices [197, 200, 201]. Non-human clinical studies that include pet dogs with cancer are designed and implemented with the humane care of the pet animal cancer patient as a primary consideration, with the informed permission of the pet owner, and under the guidance of an accredited institutional animal care and use committee (IACUC) and Clinical Trials Review Board (CTRB) [197, 200, 202]. Oversight provided by IACUC and CTRB ensures patient safety through protocol review before study initiation and monitoring patient outcome and protocol compliance during the course of a clinical trial. Additionally, the inclusion of a data safety monitoring board (DSMB) in veterinary clinical trials provides an additional layer of patient protection and assurance regarding appropriate study conduct and termination. Similar to a DSMB responsible for the oversight of a human clinical trial, the DSMB in a comparative oncology study has a role in reviewing adverse events that occur during the clinical trial and an interim analysis of trial efficacy [203]. In order to provide minimum standards for their conduct in dogs, studies are conducted using the guidelines outlined in the spirit of Good Clinical Practice (GCP). With the goal of providing assurance of credible study results, the GCP guidelines were developed by the International Conference on Harmonization to i) protect the welfare and rights of human clinical trial participants, and ii) define standards for trial design and conduct, data capture and analysis, and auditing and reporting of clinical trials [204].

Clinical translation in dogs is not constrained by the historic conventions of phase I, II and III studies. However, veterinary clinical studies contain many of the same components of human trials. For instance, informed consent and the designation of specific inclusion and exclusion criteria are included in veterinary clinical trials to ensure enrollment of proper patient populations and validate the integrity of the clinical data obtained. Central to the conduct of successful veterinary clinical trials is appropriate trial design, including adequate statistical

power to answer the specific questions that are necessary for moving product development forward. For example, assessment of biological endpoints addressing the mechanism of action or therapeutic index, identification and validation of biomarkers, and the correlation of these endpoints with imaging and pharmacokinetics [92, 205, 206]. Later studies may prioritize anti-tumor activity against measurable tumors or against minimal residual disease and should be flexible in design so as to efficiently respond to new data generated both within and outside of the study. The active pharmaceutical ingredients (API) considerations for non-human clinical studies does not require good manufacturing practice (GMP) certification, but agents prepared for these studies are prepared sterile, endotoxin free, and of high quality and purity. Given the scientific and translational intent of such studies, the use of GMP quality agents will likely become more important to study sponsors as an agent progresses to and beyond the point of investigational new drug (IND) filing [197].

Within the practice of human oncology, standardized response criteria for malignant tumors allow for more consistent and meaningful comparisons of treatment protocols and outcomes. The World Health Organization established standardized criteria for the recording of baseline data relating to the patient, the tumor, laboratory and radiologic findings, the reporting of treatment, grading of acute and subacute toxicity, reporting of response, recurrence and disease-free interval and reporting results of therapy for clinical trials in human oncology [207]. Subsequent documents establishing response criteria for malignant lymphoma and the human Response Evaluation Criteria in Solid Tumors (RECIST) have been published and are periodically updated, providing standardized criteria to assess therapeutic response that are now considered standard practice in human clinical trials [208, 209]. Given the increased number of prospective clinical trials performed in veterinary oncology and the role of spontaneous canine tumor models in the human cancer drug development process, standardized adverse event and grading criteria and guidelines to evaluate tumor response in dogs with peripheral nodal lymphoma and solid tumors have been established by the Veterinary Comparative Oncology Group (VCOG) [208, 210, 211]. These guidelines have been modeled after similar human response evaluation criteria in an effort to provide consistency and accuracy of reporting patient response. Additionally, tools to assess health-related quality-of-life in dogs with cancer have been established to assess owner-perceived changes in dogs undergoing therapy, and to evaluate

the impact agents that are administered for prolonged dosing periods have on overall quality of life [212].

Guidelines for regulatory oversight and standards for reporting data from comparative oncology trials that include pet dogs intended to support the development of human drugs are not well-defined, but specific guidelines for the timing and nature of such reports is currently under discussion. Studies in tumor-bearing dogs are typically conducted at two points in time in the life of a new human cancer treatment, either before an IND is filed (i.e. pre-IND) or after an IND is filed (i.e. post-IND); however, the implementation of the study (including protocol development and design), IACUC, CTRB and data safety management oversight, and trial conduct are similar in both development settings. The filing of an Investigative New Animal Drug (INAD) through the FDA-Center for Veterinary Medicine (CVM) is necessary for the regulatory development of a new drug for use in the animal health market, but investigational agents intended for human use alone or for those agents not yet submitted for IND status may be treated similarly to other preclinical studies in traditional model species [197].

Trial implementation guidance, including both IACUC and DSMB oversight, adequately addresses the questions of risk of the proposed studies to pet animals required for INAD filing and typically includes details on the API beyond those generally required by an INAD. Whether or not an INAD is filed for tumor-bearing dog studies conducted in the pre-IND setting, it is recommended that a full report and associated preliminary data documenting expected and unexpected adverse events should be maintained as part of the legacy of the agent under development. Unexpected adverse events that occur should be reviewed and addressed during the study by the sponsor, investigators, and DSMB. If the agent in question progresses through development, a final study report and associated data generated in tumor-bearing dog studies should be included in an IND application package. For new human cancer agents that are post-IND, data from a tumor-bearing dog study would become part of a reporting package for the agent being evaluated and regulations regarding adverse event reporting are provided by Investigational New Drug Application section 312.32 IND Safety Report [197]. Based on this criteria, any event that occurs in a study of a new human cancer agent conducted in tumor-bearing dogs that is either not serious or is expected, based on the protocol and informed consent, does not require expedited reporting. If an unexpected adverse event does occur in the setting of preclinical evaluation of a new drug in pet dogs, the inclusion of stopping rules that allow

expansion of treatment cohorts may help determine if an unexpected event is reproducible. If this is the case, modifications to the study to alter eligibility and exclusion criteria or additions to monitoring strategies may help investigators better understand these adverse events [197].

8.4.4.3

Examples of successful preclinical investigations in pet dogs with cancer

Dogs have historically been useful, informative models in the preclinical development and discovery of novel cancer therapeutic strategies. Early studies evaluating the biological features, chemotherapeutic and surgical treatment of spontaneous tumors in dogs have provided guidance for conducting similar investigations in human patients and subsequently informed the management of human cancers [65, 74, 213]. Studies in dogs are uniquely positioned to evaluate the efficacy and feasibility of novel drugs and drug delivery devices and can inform the go/no-go 'decision gate' in clinical drug development.

Examples of such efforts include comparative studies conducted in pet dogs with OS to evaluate the safety, efficacy, and feasibility of novel inhalation therapies in the management of macroscopic pulmonary OS metastasis [214, 215]. Similarities in the respiratory anatomy and tropism of OS metastasis to the lungs in humans and dogs provided rationale for early canine trials of inhaled cytokine and cytotoxic chemotherapy trials. Inhalation cytotoxic chemotherapy trials in dogs demonstrated that aerosolized therapies were well tolerated with no dose-limiting hematologic or biochemical toxicity and minimal histologic lung pathology. Furthermore, these studies supported the proposed mechanism of antitumor activity associated with this therapy in dogs and contributed to the clinical development of inhalation approaches in humans [60, 150, 214-217]. Proof of concept studies evaluating inhaled cytokine immunotherapy in dogs established a safety and efficacy profile for inhaled liposomal IL-2 therapy, demonstrated evidence of local immunomodulatory effects, and provided support for subsequent early-phase trials of this novel treatment approach in humans with pulmonary metastasis [215, 216].

Comparative studies in dogs with cancer provide an opportunity to evaluate the feasibility and tolerability of surgical procedures or devices, interventional radiology techniques, and radiotherapy that are difficult to recreate with other animal model systems [51, 74]. Early studies in pet dogs with OS undergoing various limb-sparing surgical approaches provided meaningful data relevant to limb-sparing techniques, bone allograft antigenicity, implant loosening and

fracture, and durable allograft healing in people [58, 218-220]. Similarly, clinical studies assessing image-modulated radiation therapy (IMRT) and tomotherapy technology in pet dogs with cancer provided important information regarding device utility and accuracy prior to its widespread use in human patients [221, 222]. More recent studies in dogs with nasal and bladder cancers treated with IMRT have explored the dosimetric impact of daily setup variations during treatment and dose escalation strategies; and they provided proof-of-principle that conformal avoidance radiotherapy can decrease the incidence of acute and late toxicity to surrounding normal tissues [57, 223-225].

In contrast to rodent models of cancer, dog cancers are more akin to human cancers in terms of physiology and metabolism for most organ systems and drugs. Additionally, spontaneous canine cancers occur in the presence of an intact host immune system and possess natural tumor heterogeneity, stroma, and vasculature [51, 195]. To this end, the inclusion of tumor-bearing dogs in high-resolution comparative imaging and cross-species validation studies may provide data better describing off-target tracer localization, biodistribution, and key complex intracellular processes such as hypoxia, DNA proliferation, and glucose metabolism [226-231]. For example, multimodality functional imaging studies performed in dogs with spontaneous sarcomas and carcinomas evaluated the uptake of a novel hypoxia specific radiotracer [^{61,64}Cu]copper(II) diacetyl-di(N^4-methylthiosemicarbazone) (Cu-ATSM) and (18)FDG in conjunction with paired pimonidazole hypoxia immunohistochemistry. This study demonstrated a strong positive correlation between the distribution of pimonidazole staining and uptake of (18)FDG and (64)Cu-ATSM, providing support for future clinical investigations evaluating the suitability of functional biotracers for *in vivo* imaging and radiotherapy target definition in solid tumors and their potential as radiotherapeutic agents [232]. Subsequent preclinical studies investigating intratumoral uptake of (18)FDG, (18)FLT, and Cu-ATSM in canine sinonasal tumors provided meaningful data on histology-specific positron emission tomography (PET) correlations, and on the spatiotemporal stability of tracers during radiation therapy treatments (see Fig. 8.4.6) [233, 234]. These findings may have a significant impact on clinical dose painting strategies used to define radiation therapy dose escalation.

More recently, efforts to include dogs in the preclinical modeling of personalized medicine have evaluated the feasibility of conducting molecularly-guided analysis of tumors from dogs with naturally occurring cancers in a clinically relevant setting [235]. This proof-of-

concept study demonstrated that the collection and turnaround of canine tumor samples, centralized pathology, molecular profiling and bioinformatics analysis matching gene expression to therapeutic options is achievable in a practical clinical window (<1 week). Furthermore, as observed in human personalized medicine trials, tumor gene expression signatures in dogs clustered by cancer type, whereas patient-to-patient heterogeneity resulted in drug predictions that were uniquely patient defined. This study serves as rationale for dogs with spontaneous cancers as a clinically relevant comparative model for optimizing the delivery of personalized medicine strategies and translating this to human patients.

Naturally occurring cancers in pet dogs more accurately model the complex biology of the cancer microenvironment, including antitumor immunity and angiogenesis [103, 236-242]. Experimental mouse tumor models, including transplantable models, genetically engineered tumor models, and humanized mice have provided key mechanistic insights into host antitumor immune responses that have guided the development of novel immunotherapeutic strategies in human cancer patients [243, 244]. Canine cancers, however, develop in the context of an intact immune system that recapitulates the considerable heterogeneity that exists in the human immune system. To this end, several immunotherapy trials in dogs have evaluated autologous vaccine strategies [245, 246], biological response modifiers [247-250], passive immunotherapy [251, 252], adoptive T-cell therapy [253] and have explored their feasibility and immunologic efficacy when used in combination with dose-intense chemotherapy [248, 249].

Early studies evaluating the efficacy of novel DNA-based vaccine strategies for melanoma in the setting of minimal residual disease in dogs were initially intended to inform human studies; however, these trials ultimately provided safety data required by the United States Department of Agriculture (USDA) for the approval of a veterinary biological vaccine (ONCEPT™ Canine Melanoma Vaccine, Merial), the first and only USDA-approved therapeutic vaccine for the treatment of cancer in either animals or humans [254-256]. Initial clinical evaluation of the anticancer immune effects associated with the administration of liposome-encapsulated muramyl tripeptide-phosphatidylethanolamine (L-MTP-PE) conducted in dogs with OS in the setting of minimal residual disease demonstrated single-agent anticancer activity [257]. Subsequent randomized, placebo-controlled clinical trials of L-MTP-PE conducted in conjunction with standard-of-care chemotherapy in dogs [258] served as scientific rationale for phase III evaluation of L-MTP-PE in pediatric OS. Findings of a Children's Oncology Group

clinical trial were similar to those reported in the initial canine studies and demonstrated that the addition of L-MTP-PE to standard chemotherapy in pediatric OS significantly improved overall survival at 6 years [259, 260]. Based on these findings, L-MTP-PE (mifamurtide, Mepact®) was approved for the treatment of metastatic OS by the European Medical Association. This highlights the utility of pet dogs with spontaneously occurring cancers for the investigation of novel immunotherapeutic agents, and to define the activity of these therapies in the micrometastatic disease setting.

Similarly, the development of anti-angiogenic or vascular targeted agents has been complex and with disappointing results from studies in tumor-bearing mice and human cancer patients [261]. Since neither tumor nor tumor vasculature are present in healthy animals (e.g. purpose-bred research dogs), the integration of tumor-bearing dogs in trials using anti-angiogenic agents that are under development for humans can provide a strong preclinical basis for first-in-man studies in human cancer patients. Evaluation of the targeted delivery of TNFα with an adeno-associated virus phage (RGD-A-*TNF*) to tumor and tumor-associated vascular endothelium in pet dogs with spontaneous cancer helped define the safety, selective tumor-specific localization, and efficacy of RGD-A-*TNF* targeting to tumor vasculature [262]. Serial tumor and control biopsies taken before and after the administration of RGD-A-*TN* established adeno-associated virus phage homing to disease-related vasculature and tumor-associated vascular expression of TNFα and correlated these findings with drug exposure and RECIST-based objective responses in several tumor types. Similarly, evaluation of anti-angiogenic thrombospondin-1 (TSP1) peptide mimetics (ABT-526, ABT-510) in pet dogs with cancer conducted in parallel with human clinical trials provided guidance on the single-agent drug activity, drug levels and prolonged exposure duration required for anti-tumor activity [263]. This emphasized the importance of evaluating these agents in novel trial designs and in combination with other anticancer agents. Subsequent studies have explored anti-tumor efficacy of alternative drug formulations of TSP1 peptides in dogs with soft tissue sarcomas, and the cooperative activity between cytotoxic chemotherapy and TSP1 anti-angiogenic therapy in dogs with lymphoma [264, 265].

The integration of studies that include pet dogs with cancer into the development path of new cancer drugs is becoming more common and is expected to increase as part of innovative drug development. The addition of comparative oncology studies in the preclinical setting has

the potential to improve the development path of new cancer drugs by identifying and removing drugs earlier that are less likely to succeed in human phase I trials [195, 197]. Comparative studies in dogs performed during or after phase I studies may provide meaningful data on pharmacokinetic/pharmacodynamics end-points, classifying responding patient subsets, and identifying optimal drug combinations that may eliminate inactive drugs prior to advancing to human phase II clinical trials, or optimize the design of those trials. Additionally, such studies conducted in the adjuvant or minimal residual disease setting may prioritize those agents most promising for phase III development [197, 201]. Collectively, the elimination of inferior drugs earlier in the drug development pathway will reduce the number of drugs entering later phases of drug development and increase the overall success rate of phase III trials. This will substantially decrease the risks and costs of drug development.

To this end, studies in dogs with cancer can aid in establishing relationships between a cancer target, its modulation with a novel small-molecule inhibitor, and clinical benefit. Correlative studies that would be challenging to conduct in humans, including multiple biopsy and other biospecimen collection time points, are readily feasible in pet dog studies. A prospective dose escalation study of rapamycin in dogs with OS was performed to define optimal dosing schedules, assess biomarkers, and provide the rationale for the use of rapamycin, or potentially other mTOR inhibitors, in OS [198]. Pre- and post-treatment biopsies and peripheral blood mononuclear cells (PBMC) were collected and 48-hour whole blood sampling was performed to establish a pharmacokinetically-relevant and pharmacodynamically active dose of rapamycin in dogs with OS. Data from this phase I trial demonstrated that biologically effective concentrations of rapamycin were safely obtainable in dogs and provided evidence of target modulation in tumor tissues and PBMCs. This study highlights the advantage of integrating the comparative approach in the development path of new cancer drugs. Importantly, such studies help to establish critical pharmacokinetic and pharmacodynamic relationships so that drugs with an unfavorable therapeutic index or inferior target modulation attributes may be identified and removed from development earlier in the process [195, 196]. That thus improves the identification of the agents most likely to succeed in human clinical trials.

This comparative approach also can have a significant impact on informing the development and conduct of later-stage studies in humans. Heat shock protein 90 (HSP90), a molecular chaperone that promotes the conformational maturation and stabilization of a wide

variety of client proteins, is a promising target for therapeutic intervention in cancer. Ganetespib, a novel small molecule inhibitor of HSP90, and its water soluble prodrug STA-1474 demonstrated activity against canine OS and malignant mast cell lines *in vitro* and induced tumor regression, apoptosis and downregulation of key targets including MET and AKT in OS xenografts [266, 267]. Based on these findings, a phase I study of STA-1474 was performed in dogs with spontaneous tumors [205]. This clinical trial demonstrated biological activity in various canine cancers (including dogs with metastatic OS and malignant mast cell disease), established safety and toxicity profiles, and provided important information regarding expected gastrointestinal adverse events that were subsequently observed in human clinical trials with ganetespib. Pharmacodynamic endpoints provided evidence of target modulation in PBMCs and established this as a reliable biomarker of drug activity. Additionally, pharmacokinetic analysis in this study provided information on drug levels and exposure duration that subsequently established dosing schemes that were unanticipated prior to the start of the clinical trial. These data laid the groundwork for the current clinical evaluation of ganetespib in humans.

Additional examples exist that highlight the use of dog models to inform preclinical drug development, including agents targeting pathways associated with invasion and metastasis [268]. Large animal spontaneous cancer models compliment the use of both existing rodent models and human clinical trials. Their inclusion in preclinical and translational studies will facilitate the rapid intermediate evaluation of agents prior to or after early human trials. This will result in a more informed and optimal cancer drug development pathway.

8.4.5

Necessary developments for realizing the potential of canine models

8.4.5.1

Epidemiology, longitudinal cohorts, tissue repositories and integrative genomics

Great efforts have been made to put canine models on equal footing with other investigational animal models. Much of that work has focused on creation of genetic resources, validation of human relevance and development of resources for clinical trials [9, 195, 196]. Here we note several areas that are likely to have a high impact on biomedical canine models [7]. First, the gap in epidemiology of common diseases in dogs should be addressed. Specifically, longitudinal data in veterinary registries is necessary for both epidemiological and translational research [195, 196,

269]. There are many examples of early efforts in this [270-273, see p. 55 of 274]. In terms of utility for understanding the common diseases of complex genetics, such resources could stress this aspect in dogs and cats. The information would be most useful if it was, at least in part, associated with biological samples. That could lead to powerful databases for integrative analysis of clinical, environmental and molecular (genetic, epigenetic and other 'omics) data in select breeds. One major limitation to these goals is the lack of standardized veterinary records. The creation of national or international standards, and ideally a common electronic veterinary medical record system, would be major advances. While there are individual academic and national veterinary tissue banks (http:// .ccogc.net/), these could be developed further and combined with electronic veterinary health records. It is generally challenging to acquire complete dog pedigree data, but this can be invaluable in genetic studies. An ideal longitudinal cohort should make it possible to gather the following data in electronic form: clinical, environmental, pedigree, genetic and molecular phenotypes.

8.4.5.2

Improved genome annotation and development of key research areas

The quality and completeness of the canine genome assembly is among the highest of any mammalian genome sequenced to date [23, 275]. However, compared to the human genome annotation, the dog's is in its infancy. Many dog genes are not annotated, but can only be found through mapping of genes from other species. This often requires user verification to determine gene identities, gene structure, etc.; and in many cases there is little or no information on mRNA expression and splicing in dogs. Some of the necessary improvements could be done computationally, but that would require significant manual effort and would still require experimental confirmation. The most effective approach may thus be to conduct multidimensional genomic surveys as were done for SNPs (to create the most recent SNP platform [276]) or breed-specific CNV maps [40, 41]. The use of canine models could be thus be accelerated by expanding genomic annotation for individual breeds of high interest. Among the data that would be most useful are haplotype maps and whole genome sequences. Other high-value information is tissue-specific genomic data, beginning with RNA expression (underway; [275]) and DNA methylation (and in the future adding other genomic dimensions as is being done in humans [277]).

Because dogs are not commonly used as experimental models in the way that, say, zebrafish, frogs, chicken and mice are, their molecular genetics/biology are not well developed. Going forward, it will be important to strengthen these aspects in order to capitalize on the rich genetic findings. For instance, the alpha and beta globin gene clusters that encode the subunits of hemoglobin are among the best-characterized gene-cluster regions across vertebrates and mammals. However, those genes are not annotated in the most recent dog genome assembly, and they cannot be readily discerned by visualizing the human genes that map by homology. This means that investigators interested in comparative blood studies in dogs either have to ignore hemoglobin gene regulation and protein sequences (or study it blindly without that information) or characterize the gene structure, protein sequences, and gene cluster regulation before undertaking those studies. Some of the most important areas that can dramatically raise the value of dog models are immunity, metagenomics, metabolism, thyroid biology, obesity, aging and stem cell biology. Of these, the one where dogs may have the most promise is aging. Because of the dramatic range in longevity across breeds, dogs may represent the best animal in which to study aging [278]. Moreover, canine longevity correlates inversely with size across dog breeds and has clear associations with reproductive traits and mortality/cause of death (including infection and cancer) [170, 279-282]. Dogs may thus offer the best opportunity to dissect apart the genetics of growth and aging, and to elucidate their relationships to longevity and mortality/cause of death (including from cancer). With regards to cancer, it will be important to expand research in areas that include metastasis, tumor microenvironment, and resistance to chemotherapy. Lastly, there are several technological aspects of canine models that could be further developed to enable their common use in translational veterinary studies. These include gene therapy, antisense nucleic acids (which can be used to knockdown gene expression, for exon-skipping, or for blocking specific long non-coding RNAs to, in effect, increase gene expression), genome editing, and stem cell therapies.

8.4.5.3

Opportunities for understanding the complete biology of spontaneous cancers

Comparative oncology approaches open unique and unparalleled opportunities to potentially identify targets for therapeutic intervention in canine and human cancer. This effort will require increased rates of canine tumor archiving and genetic-profiling [275, 283]. Many examples

demonstrate how canine genetics can reveal cancer subtypes associated with clinical outcomes in dog and human patients [52, 79-81, 106]. To translate these findings more globally into the cancer drug development pathway, the establishment and molecular characterization of canine tumor cell lines from a variety of tumor types is needed to validate molecular target expression and provide preliminary assessment of the biological effects of therapeutic agents in vitro. Data validating the expression of molecular targets in primary canine tumor samples will provide rationale for potential therapeutic benefit and demonstration of biological activity of a therapeutic agent against tumor cell lines will support the subsequent in vivo evaluation of novel cancer agents in dogs with spontaneous cancers.

Spontaneous cancers in dogs model the biological complexity of human cancers, including the complex interaction of the immune system with tumor cells and the development of chemotherapeutic resistance and metastasis that is integral to outcomes in both humans and dogs. For example, naturally occurring OS in dogs serves as a useful large animal model to study the biology of metastatic progression and identify potential targets that are linked to the pathogenesis of micrometastatic progression [72, 201]. Tissues from metastatic lesions and matched primary tumors from human patients are not widely available at this time; however, expanding existing biospecimen efforts to collect clinically annotated tissues from dogs with OS throughout the course of disease presentation and progression may be feasible given the high incidence and aggressive metastatic behavior of this disease in dogs. This effort may provide a broader understanding of the development of metastasis in OS through candidate drug target expression profiles and their dynamics during metastatic progression. Similarly, the identification of populations of cell that have the ability to suppress antitumor immune responses have been challenging to dissect in animal models; however, components of the tumor microenvironment, including regulatory T cells (Tregs) and myeloid-derived suppressor cells have been characterized in healthy and cancer-bearing dogs [239, 284, 285]. Furthermore, the clinical significance of various components of the tumor microenvironment has been recognized in several cancer types in dogs, providing additional support for the notion that systemic antitumor activity plays an important role in the pathogenesis of naturally occurring cancers in dogs [103, 242, 286]. The development of spontaneous cancers in dogs provides the opportunity to study tumor progression in the context of an intact immune system where tumor, host, and tumor microenvironment are interacting intimately with one another. To this end, studies evaluating

metronomic chemotherapy strategies in dogs to halt or slow tumor progression may help define their proposed mechanisms of action, including anti-angiogenic or antitumor immune responses and identify pharmacodynamics markers of effective therapeutic exposure and target modulation in tumor and surrogate tissues [240, 241].

8.4.5.4

Development of high-impact programs in preclinical cancer studies

Historically, the widespread use and integration of comparative studies was hindered by the lack of infrastructure to coordinate the necessary parties: veterinary and human oncology, drug industry and centralized biospecimen repositories. The organization of collective groups (including the COTC, CCOGC, and Veterinary Cooperative Oncology Group) now provides the infrastructure and resources to support studies of pet animals with cancer that are increasingly "integrated" into the developmental path of new cancer treatments. For these studies to be most useful, they must be designed so as to ask important and well-defined biological questions that cannot be easily or fully answered by conventional preclinical models or in early human clinical trials [197]. The inclusion of a multi-disciplinary collaborative team (including basic scientists, veterinarians, and physicians) to rigorously review such investigations is necessary to ensure that questions are prioritized and efficiently answered. To optimize the translation of findings in the canine model, preclinical studies conducted by the COTC involve a multicenter collaborative network of veterinary teaching hospitals to ensure that they are adequately powered with sample sizes sufficient to detect drug effects with the endpoint parameters used. Incorporation of standardized response criteria (RECIST) and adverse event reporting in preclinical trials has facilitated the comparison of current and future treatment protocols used for companion dogs with cancer. Novel electronic reporting systems designed to provide real-time data capture are being successfully used in canine studies and it is anticipated that this method of data reporting may be incorporated into larger, multi-institute investigations to encourage discussion among study investigators during all phases of the investigation [51]. Importantly, translational studies are optimally designed and developed to include study endpoints that are specifically aligned with the design and development of studies in other preclinical species and human studies, thereby expanding the totality of information available for an agent and complementing data from other preclinical models that will optimize the drug development pathway.

As part of the broader field of comparative oncology, translational drug development studies in dogs with cancer have classically evaluated pharmacokinetic and pharmacodynamic endpoints to define dose and schedule for therapeutic agents. Preclinical studies in dogs evaluating novel small molecule inhibitors have established relationships between drug exposure and target modulation; however, changes in target expression within peripheral blood mononuclear cells following drug treatment have aided in the discovery of surrogate biomarkers that may predict biological response [198, 205]. To complement this effort, the expansion and integration of molecular imaging techniques in tumor bearing dogs may provide critical new data regarding surrogate imaging endpoints during the process of drug discovery and development. Comparative imaging studies in dogs with cancer have established proof-of-concept for radiolabeled tumor-targeted peptides, and provided validation data for novel PET radiopharmaceutical agents, and have demonstrated the feasibility of serial PET/CT imaging to monitor response to cytotoxic chemotherapy or small molecule inhibitors [226, 287-290]. Diagnostic imaging studies in dogs have largely focused on improved methods for early determination of response to therapy or prognostication; however, alternative imaging modalities such as optical imaging of tumor tissue beds to detect minimal residual disease at the time of surgery have been explored in dogs with solid tumors to inform their development for humans [291, 292]. The incorporation of dogs in such comparative imaging studies designed to validate and optimize novel imaging techniques will be critical to imaging agent development and in the drug development pathway.

Lastly, the opportunity to conduct studies in dogs in the setting of microscopic or micrometastatic disease is a unique and powerful advantage of the canine model [51, 196, 258, 293, 294]. The advantage of including pet dogs with cancer in the development of drugs that target metastatic progression in OS has been recognized and more recently, consensus guidelines to establish a preclinical translational pathway to value and prioritize potential therapeutic agents uniquely targeting the metastatic phenotype have been proposed [201]. A major effort of comparative oncology has been to identify shared targets in human and canine OS. Substantial progress has been made through experimental and preclinical investigations to identify dysregulated intracellular signaling pathways likely contributing to OS metastasis. For example, data from in vitro and murine xenograft studies have identified signal transducer and activator of transcription 3 (STAT3) and the membrane-cytoskeletal linker ezrin as relevant targets for

therapeutic intervention in human and canine OS and support the current development of small molecule inhibitors of STAT3 and ezrin for the treatment of OS [83, 165, 295-297]. The inclusion of dogs with spontaneously occurring OS in the development and testing of novel therapeutics that target STAT3 and ezrin biology in the micrometastatic disease setting will likely provide important new information regarding efficacy and optimization of different doses and schedules with direct relevance to future testing in people. Importantly, while druggable targets linked to metastatic progression of cancer have been identified, many of these targets and associated processes seem to influence the progression of metastatic cells from microscopic disease to that of grossly detectable lesions. Therefore, it is unlikely that many of these agents will have a measurable effect on established and grossly detectable lesions at either the primary or metastatic location [201]. As a subset of these drugs that target metastatic may have no activity in the setting of measurable disease, preclinical trials conducted in dogs in the micrometastatic disease setting may provide compelling data defining their activity and support their advancement in the drug development pathway.

8.4.6

Key challenges and recommendations for using canine models

8.4.6.1

Challenges of population structure in dog models

The many advantages of the canine model have been discussed in this chapter, but one should also be aware of its potential challenges. Most notably, the very thing that makes dogs so genetically useful – isolated breeding – also presents limitations and complications. Breeding of domesticated dogs is not natural or random. Rather, dogs are bred for diverse types of selection by humans. This must be accounted for in dog research, especially in genetics. For instance, statistical power is affected by the extent of the difference that the design expects to detect. Keeping other variables that influence statistical power the same, a larger effect-difference is associated with increased power (i.e., that translates to an increased chance of a successful true positive detection). The larger effect sizes often observed for dog traits (compared to humans) are powerful, but this is in exchange for complex structural patterns and relatedness that can confound or mask the detection of that effect. Equally important, dogs can be fixed for the gene alleles that have the greatest effect on the overall disease/trait of interest; and this information

would be missed by genetic mapping. Population structure can generate false positives and be difficult to separate from the traits of interest. It is also possible that some traits are inextricably linked to population structure.

These issues have different implications for single-breed and cross-breed studies. In single breed studies, one must measure the effects of population structure on the GWA analysis and should identify the portion of the genome that was not sampled due to high levels of homozygosity in that breed. To address the latter with conventional methods (mainly without breeding studies), one must move laterally, first, to related breeds (likely to share disease-associated variants) and then to increasingly unrelated breeds. To address these issues across breeds, one must be careful to understand the difference between stereotypes associated with most or all members of a breed [35, 276, 298] and phenotypes that clearly segregate within breeds [42, 299-301]. If there is a mixture of breeds studied in a single analysis there should be sufficient breed variation while avoiding over-representation of any breed groups (unless they were evenly distributed across case-control groups or QTL phenotypes). That would serve to avoid strong population structure, but its effects should still be measured. If one had sufficient power, such analyses could be done in multiple GWA sets (each would have its own recognizable population structure and be analyzed separately). In the next section we propose additional ways to mitigate these challenges and suggest how investigators can use dog populations to avoid biases that are common in human biomedical research.

8.4.6.2

Recommendations for optimal results in canine preclinical research

While the previous section suggests how dog breeding practices can introduce a major type of bias – population structure – we argue that dogs offer ways of circumventing biases that can be difficult to avoid in humans. First, human studies are extremely susceptible to very strong socioeconomic impacts [302]. And second, the availability of hundreds of popular dog breeds that exist in a clinical context makes it possible to avoid the bias present in single human ethnic groups, single rodent strains and single dog breeds. This advantage can be appreciated by quickly reviewing the ways that bias is generally addressed [303]. For experimental studies, one can begin by randomizing assignment to comparison groups at baseline. In observational studies, one can begin by minimizing or stratifying heterogeneity to reduce differences between groups. And

for both of those study designs, one can then i) measure and report the groups' baseline data, and ii) if groups are unequal, describe the extent and direction of bias and discuss the effects of those on the interpretation of the study. For example, the issue of reducing differences between groups in observational studies is illustrated by the study of OS risk in three separate breed-specific GWAS's [42].

In the Introduction to this book, we discussed the causes of irreproducibility in preclinical research. Among the ways to prevent the most common problems are proper experimental (e.g., blinding and randomization) and analytical design, good record keeping, and careful reporting. Below we list key recommendations to increase the reproducibility of preclinical canine studies and to optimize their translation to human cancer patients:

- Improved genome annotation and development of breed-specific catalogs of all genetic variation for select breeds
- Improved genomic annotation of banked canine tumor biospecimens to aid in cross-species genomic analysis
- Establishment of *in vitro* canine tumor cell lines to evaluate biological activity and provide support for *in vivo* evaluation of novel therapeutics in dogs with spontaneous cancers
- Increased knowledge of key biology (e.g., immunity and aging)
- Development of centers of excellence for canine and feline common-disease epidemiology
- Development of centers of excellence for targeted biological therapies (e.g., antisense nucleic acids, siRNA, stem cells, gene therapy and genomic editing)
- Prospective, adequately powered studies with sample sizes sufficient to detect drug effects with the outcome parameters used
- Establishment and validation of primary and secondary endpoints in preclinical trials, including, but not limited to, determination of pharmacokinetic and pharmacodynamics relationships, surrogate biomarker identification, and comparative imaging endpoints
- Standardized data reporting and inclusion of electronic data capture methods; standardized methods to evaluate response criteria and health-related quality-of-life assessment
- Inclusion of dogs with spontaneous cancers in preclinical trials conducted in the setting of microscopic or minimal residual disease to inform the development of drugs targeting metastatic progression
- Increased investigation of clinical imaging methods

8.4.7

Conclusions

Going forward, the greatest contributions of canine models are likely to be in two areas. The first area is genetics – making use of their tractability for genetic mapping of complex disease risk. That will in turn lead to an unprecedented understanding of gene-gene and gene-environment effects. Predictive risk models will be built and tested in prospective studies. The second area of great promise is the full spectrum of translational medicine. This will take advantage of the large population of dogs and the established setting of veterinary health care. Knowledge of germline-risk genetics will lead to new-targeted approaches in both preventive and therapeutic medicine. For many diseases shared by humans and dogs, drugs under development for humans may be tested in veterinary clinical trials (even before they are approved by the FDA for human trials). Canine genomics and translation are converging with synergistic effects [10]. Cancer is a particularly attractive area for canine translational genomics because many targeted pathways can be assayed in tumors and that information can be used for drug selection. We propose that the strengths of dog models result in dramatically increased power and, potentially, reduced bias. Canine models thus represent a unique resource for significantly improving the reproducibility of human preclinical research.

Figure Legends

Figure 8.4.1: Popular Breeds and the Percent of Deaths from Cancer. On the Y-axis, breeds are ordered according to percent of deaths due to cancer [data from 63]. Next to breed name on X-axis in parenthesis is the popularity of breed rank-ordered for 2013 according to the AKC (https://www.akc.org/reg/dogreg_stats.cfm). All pictures used from (www.akc.org/breeds/).

Figure 8.4.2: Estimated New Cases and Deaths for Common Cancers in the USA. According to Centers for Disease Control (CDC), cancer is narrowly second only to heart disease for the leading causes of death. In 2011, heart disease claimed 596,577 lives, while cancer took only slightly less at 576,691 lives (http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm). Coupled with loss of life is also the fininacial burdan. In 2010, cancer prevelance cost have been estimated between 124.5-216.6 Billion USD [http://www.cancer.org/cancer/cancerbasics/economic-impact-of-cancer; 304]. This suggests that landmark revolutions are needed in treatment and care of cancer patients. No animal model completely recapiltureles the humans cancers perfectly, but the naturally occuring cancers in dogs more closely resemble the disease then do other animal models. Data for figure taken from *"How Common is Cancer"*, http://seer.cancer.gov/statfacts/html/all.html.

Figure 8.4.3: Number of Publications Related to Dogs and Cancer. We performed a search using the PubMed database (http://www.ncbi.nlm.nih.gov/pubmed) for publications related to dogs and cancer. We used the following search terms: "dog OR dogs OR canine OR dogs AND cancer". Years were grouped and average publications for years calculated. From this, we calculated the annual growth rate and total percent change (Percent Growth Rate = Percent Change / Number of Years; http://www.miniwebtool.com/percent-growth-ratecalculator/?present_value=391&future_value =773&num=12).

Figure 8.4.4: Germinal Centre of DLBCL. Antigen-activated B-cells differentiate into centroblasts that undergo clonal expansion in the dark zone of the germinal centre. During proliferation, the process of somatic hypermutation introduces base-pair changes that can lead to changes in the amino-acid sequence. Centroblasts then differentiate into centrocytes and move to the light zone, where the modified antigen receptor, with help from other immune cells, is selected for improved binding to the immunizing antigen. Newly generated centrocytes that produce an unfavorable antibody are removed. Cycling of centroblasts and centrocytes between dark and light zones appears to be mediated by a chemokine gradient. Antigen-selected centrocytes eventually differentiate into memory B cells or plasma cells. Centroblasts with genetic alterations that do not undergo apoptosis as expected can become GBC DLBCL. Likewise, plasma cells can become ABC DLBCL. Not Shown: Thymic cell which leads to PMBL. Listed are several known malignant transformations, adapted from [117, 305, 306].

Figure 8.4.5. Integration of pet dogs with cancer into translational drug development studies. Canine cancer models compliment the use of both conventional preclinical models (mouse, research-bred dog and non-human primate) and human clinical trials and their inclusion in preclinical and translational studies will facilitate the rapid intermediate evaluation of agents prior to or after early human trials. Translational drug development studies in dogs may answer important questions about a new drug candidate such as toxicity, biological activity, and establish pharmacokinetic pharmacodynamic relationships for an agent before it enters human studies. Importantly, the comparative approach may answer questions that emerge in early phase human trials such as optimized dosing schedules, combination therapies, and the establishment of surrogate biomarkers or molecular imaging endpoints that will inform the evaluation of these agents as they move into later stages of development. Importantly, the totality of information generated from this comparative and integrative approach will likely reduce the late attrition rate of new cancer therapeutics and contribute to the identification of agents most likely to succeed in human clinical trials. Reprinted from [51] with permission from Macmillan Publishers Ltd: Nature Rev Cancer, copyright 2008.

Figure 8.4.6. Comparative imaging study in dogs with sinonasal tumors investigating the spatiotemporal stability of Copper(II)-diacetyl-bis(N4-methylthiosemicarbazone) (Cu-ATSM) and 3'-deoxy-3'-^{18}F-fluorothymidine (FLT) positron emission tomography distributions in during radiation therapy. (Upper panel) Sagittal positron emission tomography/computed tomography (PET/CT) slices are shown from a dog with sinonasal carcinoma pretreatment (pre) and midtreatment (mid) with intensity modulated radiation therapy. Cu-ATSM (middle) and FLT (bottom) scans demonstrate stable spatial distributions of both radiotracers during therapy. (Lower panel) Voxel-based scatter plots comparing pretreatment (pre) and midtreatment (mid) Cu-ATSM and FLT standardized uptake value (SUV) distributions and Spearman rank correlation coefficients (upper left) for dogs with nasal carcinoma or sarcoma. Spatial distributions and uptake of dose painting targets Cu-ATSM and FLT remain stable through mid-treatment, regardless of histology. Reprinted from [234] with permission from Elsevier.

Figures

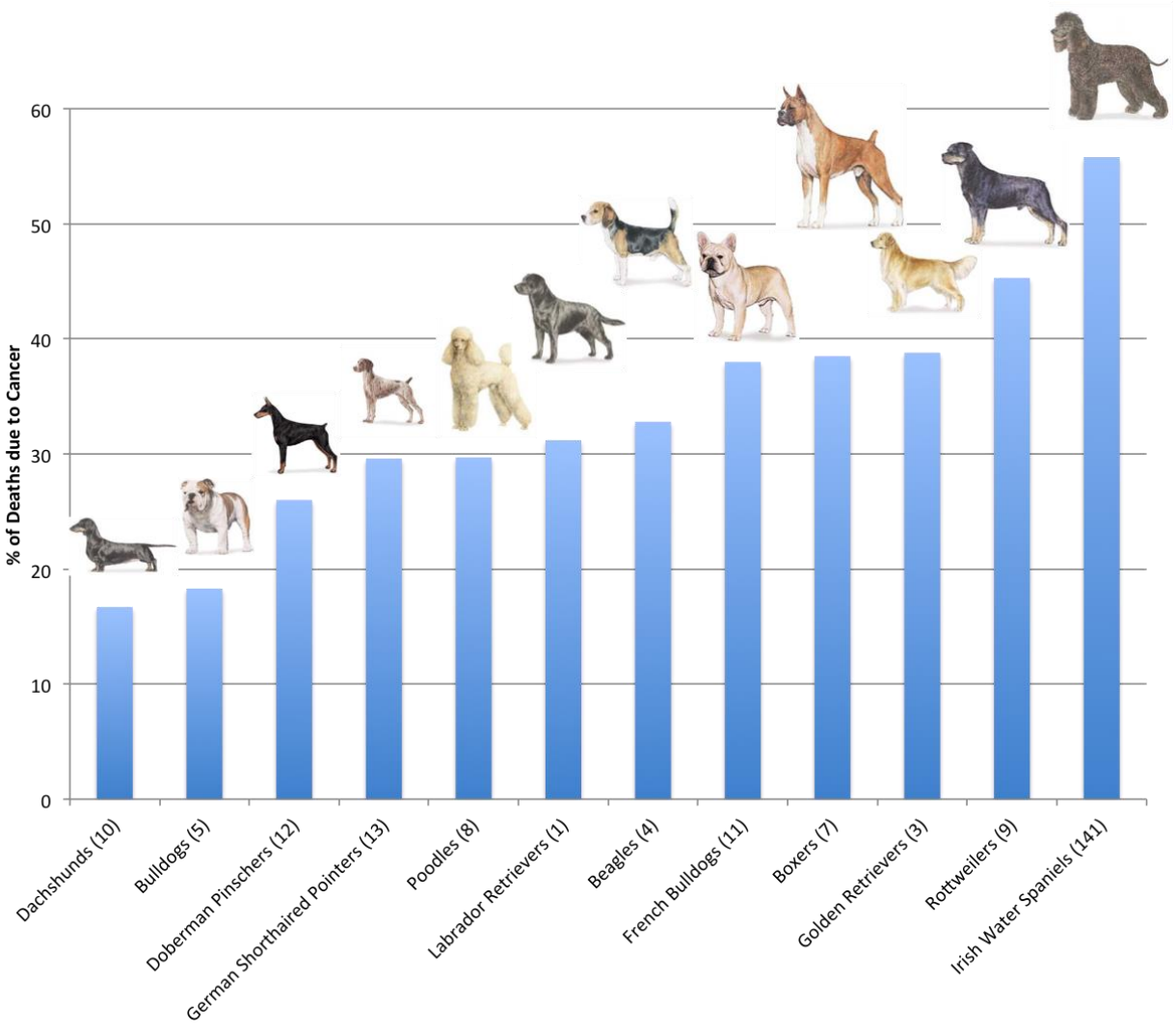

% of Deaths due to Cancer
0
10
20
30
40
50
60
Dachshunds (10)
Bulldogs (5)
Doberman Pinschers (12)
German Shorthaired Pointers (13)
Poodles (8)
Labrador Retrievers (1)
Beagles (4)
French Bulldogs (11)
Boxers (7)
Golden Retrievers (3)
Rottweilers (9)
Irish Water Spaniels (141)

Figure 8.4.1

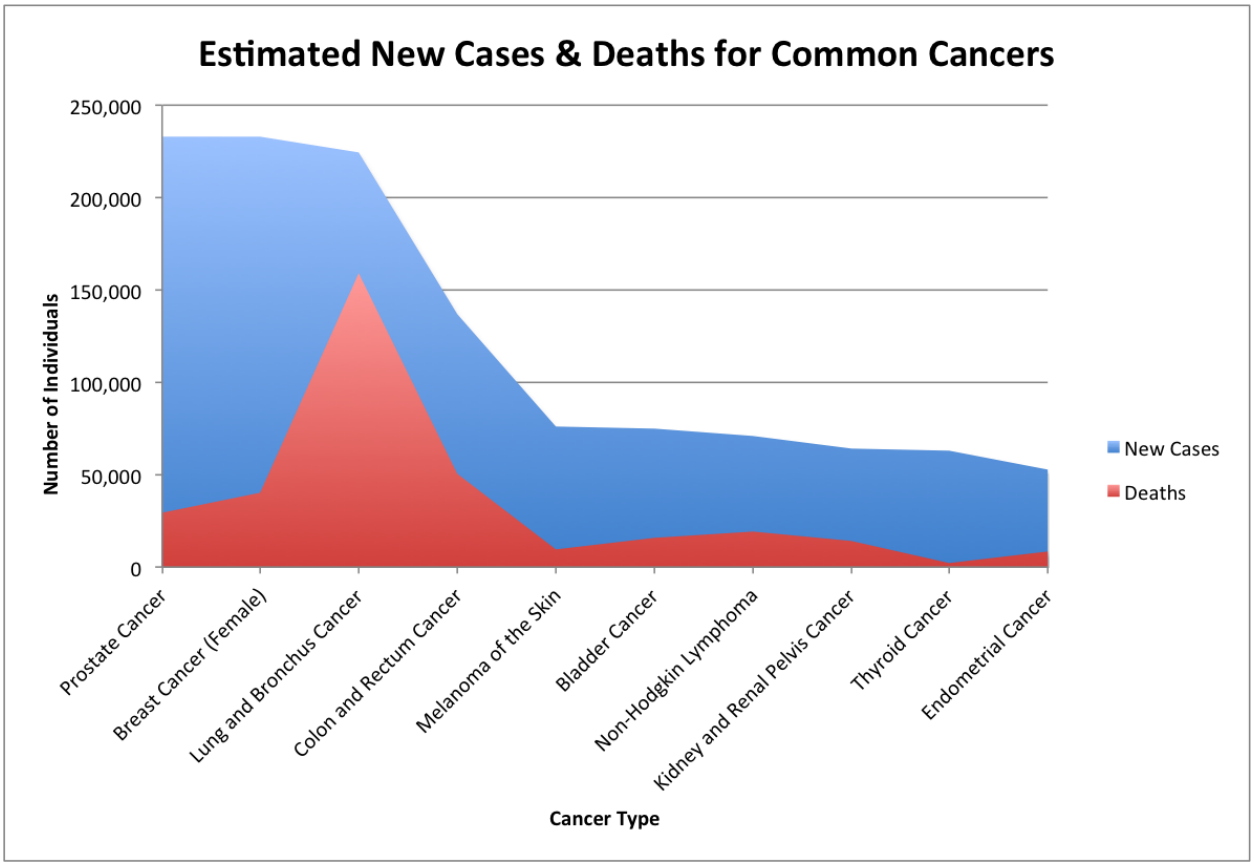

Estimated New Cases & Deaths for Common Cancers
250,000
200,000
150,000
100,000
50,000
0
Number of Individuals
Prostate Cancer
Breast Cancer (Female)
Lung and Bronchus Cancer
Colon and Rectum Cancer
Melanoma of the Skin
Bladder Cancer
Non-Hodgkin Lymphoma
Kidney and Renal Pelvis Cancer
Thyroid Cancer
Endometrial Cancer
Cancer Type
New Cases
Deaths

Figure 8.4.2

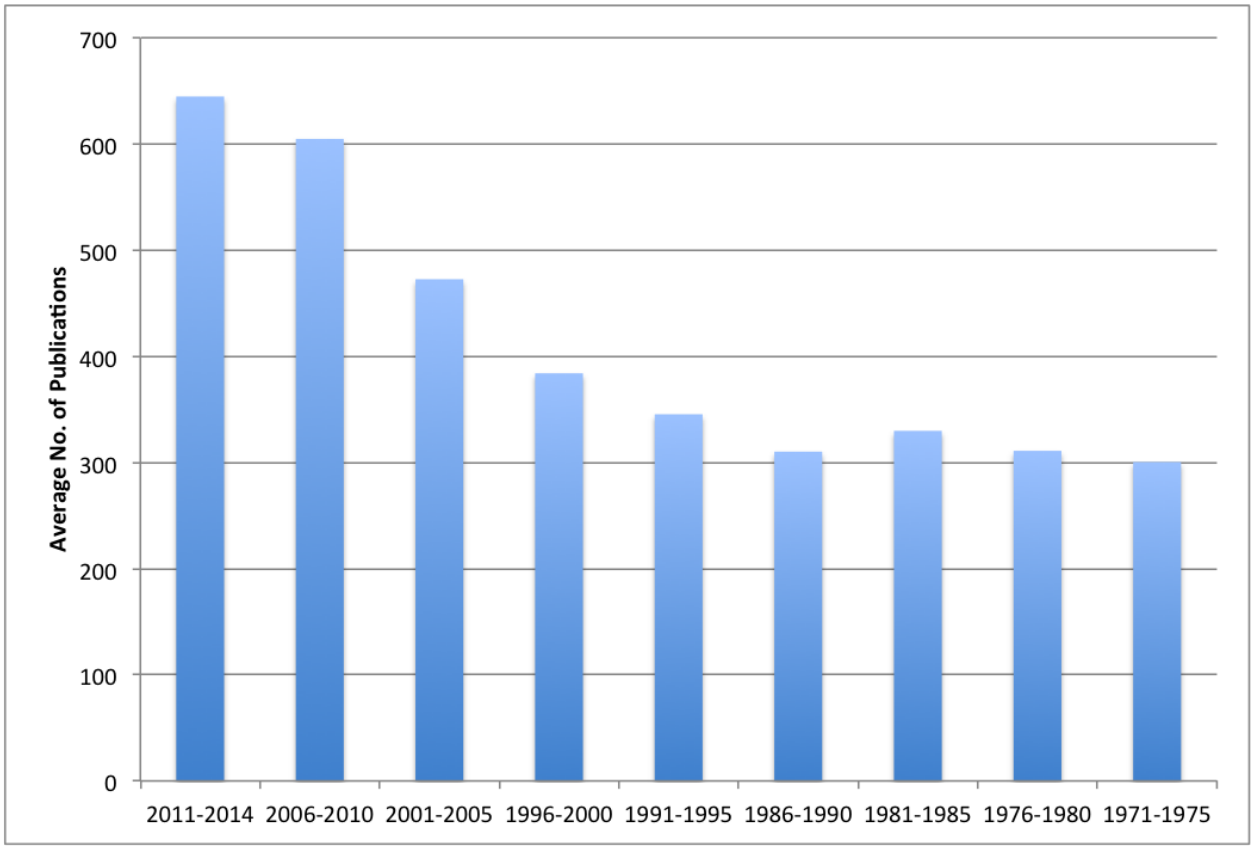

Average No. of Publications
700
600
500
400
300
200
100
0
2011-2014
2006-2010
2001-2005
1996-2000
1991-1995
1986-1990
1981-1985
1976-1980
1971-1975

Figure 8.4.3

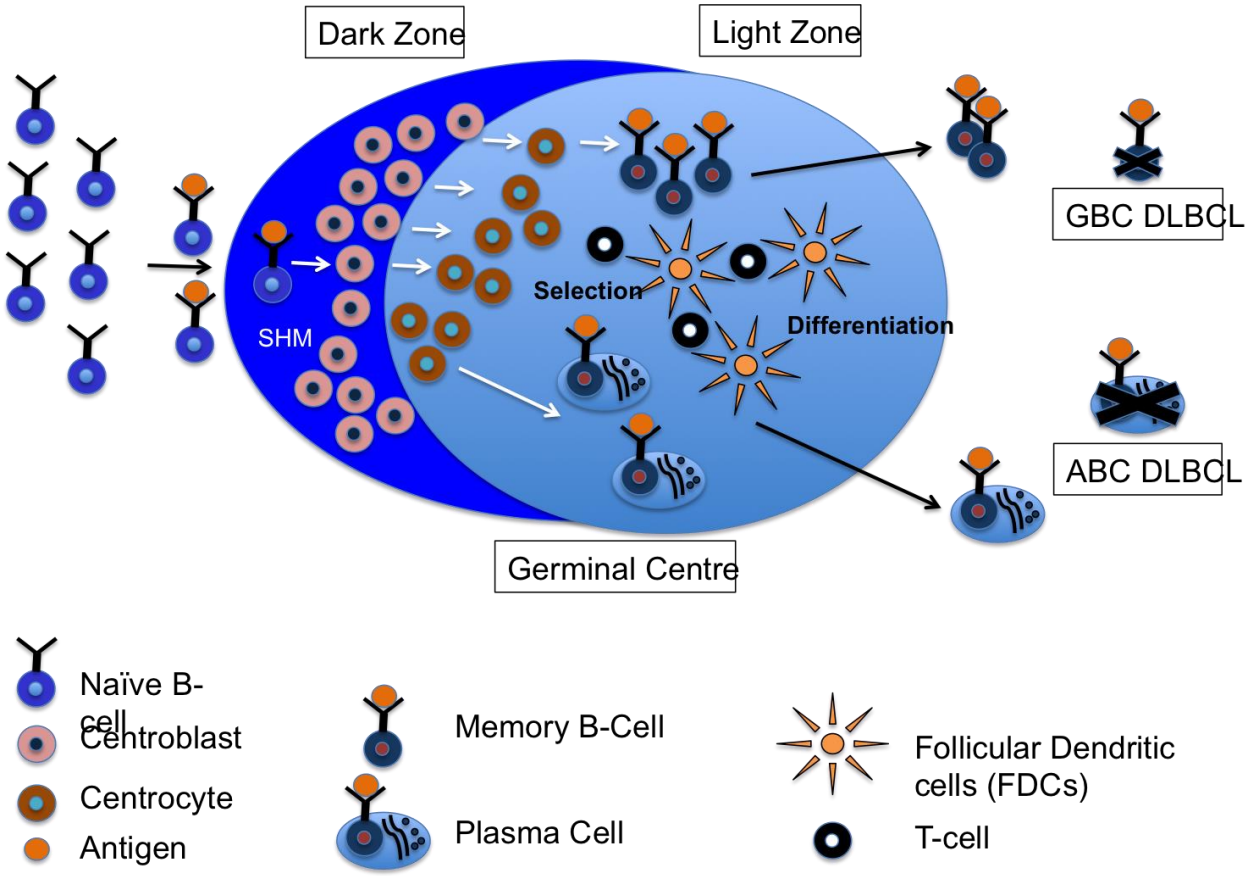

Dark Zone
Light Zone
SHM
Selection
Differentiation
GBC DLBCL
ABC DLBCL
Germinal Centre
Naïve B-cell
Centroblast
Centrocyte
Antigen
Memory B-Cell
Plasma Cell
Follicular Dendritic cells (FDCs)
T-cell

Figure 8.4.4

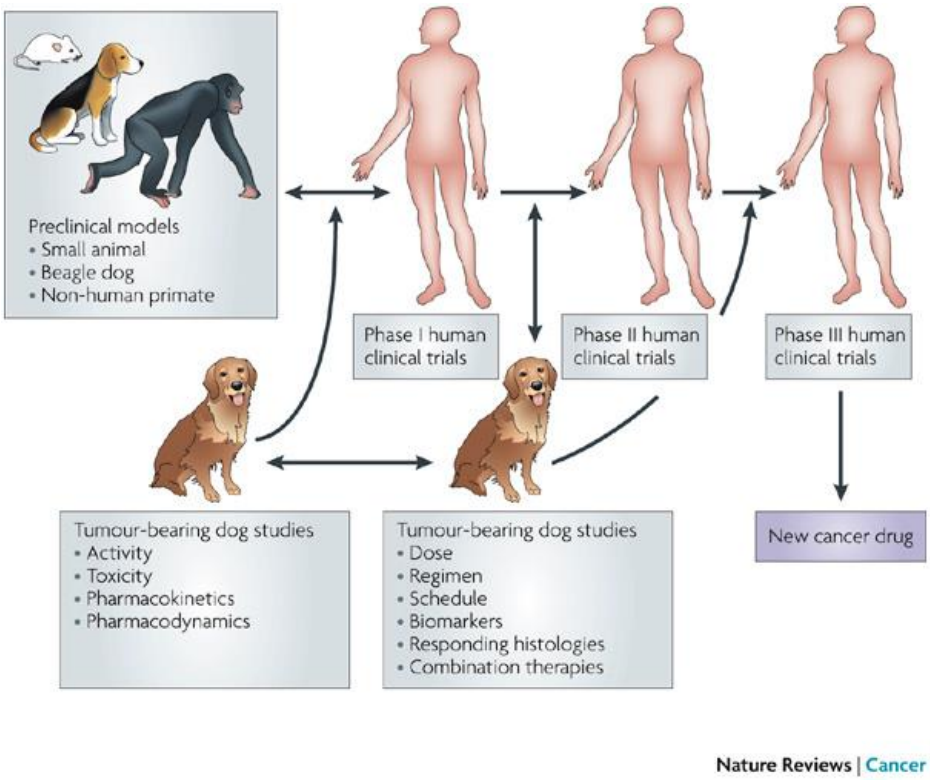
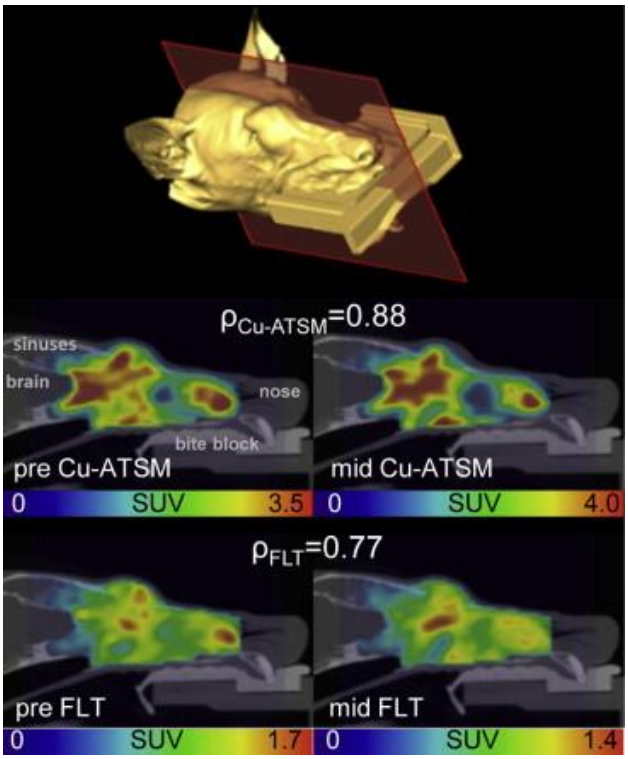

Preclinical models
• Small animal
• Beagle dog
• Non-human primate
Phase I human clinical trials
Phase II human clinical trials
Phase III human clinical trials
Tumour-bearing dog studies
• Activity
• Toxicity
• Pharmacokinetics
• Pharmacodynamics
Tumour-bearing dog studies
• Dose
• Regimen
• Schedule
• Biomarkers
• Responding histologies
• Combination therapies
New cancer drug
Nature Reviews | Cancer

Figure 8.4.5


ρCu-ATSM=0.88
sinuses
brain
nose
bite block
pre Cu-ATSM
mid Cu-ATSM
0 SUV 3.5
0 SUV 4.0
ρFLT=0.77
pre FLT
mid FLT
0 SUV 1.7
0 SUV 1.4

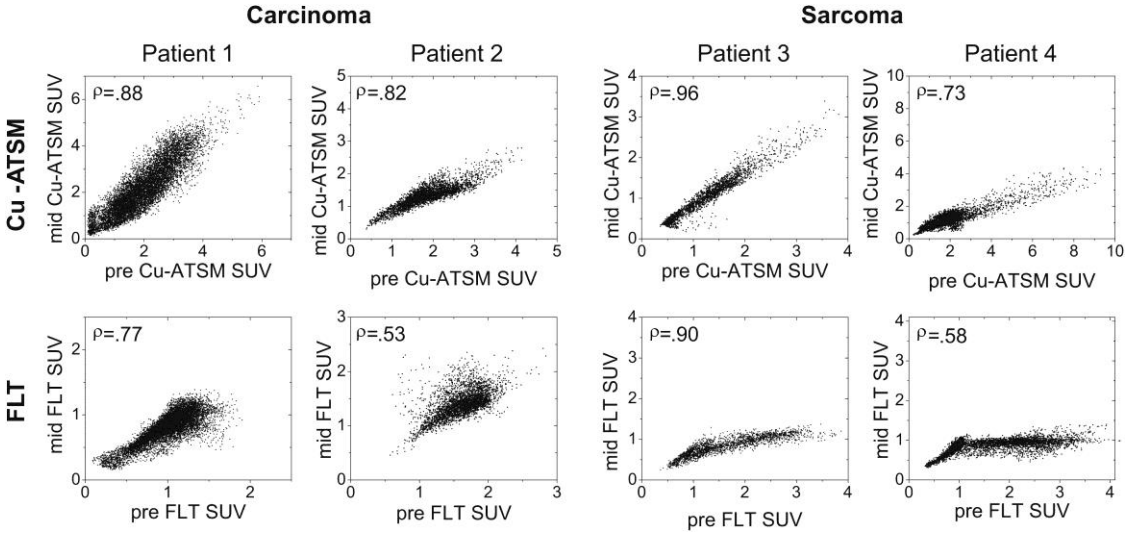

Carcinoma
Sarcoma
Patient 1
Patient 2
Patient 3
Patient 4
Cu -ATSM
FLT
ρ=.88
ρ=.82
ρ=.96
ρ=.73
ρ=.77
ρ=.53
ρ=.90
ρ=.58
mid Cu-ATSM SUV
pre Cu-ATSM SUV
mid FLT SUV
pre FLT SUV

Figure 8.4.6

Tables

Table 8.4.1 Proportional Mortality due to tumors/neoplasm over Years by Breed. Due to the genetic consequences of pure-breeding, many dog breeds have inherited very high incidence rates of specific cancer types. Here, we look at the proportional mortality for the development of a neoplasm in different breeds during different years as reported. Proportional Mortality calculated as number or deaths due to tumors and/or neoplasm /total deaths in breed x 100. Data taken from [63, 170, 307-309].

Breed	**1997** [307]	**2003** [309]	**2005** [308]	**2010** [63]	**2011** [170]
Bernese Mountain dog	32.6	34.2	41	45.7	
Boxer	29.6	11.6	37	38.5	44.3
Cocker Spaniel	22.2	14.6		29.4	27.7
Dachshund		17.3	8	16.7	8.9
Doberman	22.2	12.9	23	26	26
German Shepherd dog	14.8	9.6	16		
Golden retriever		20.3	30	38.8	49.9
Great Dane	12.3		13		22
Greyhound	12.3		11		21.6
Irish wolfhound	24.8		22		31.8
Labrador retriever		13.1	21	31.2	34
Newfoundland	16.8	5.6	14	27.1	1.99
Poodle		8.5-16.3	10		1.1-2.7
St. Bernard	13.1	5.7	19		26.9
Total Dogs Investigated	222,000	2,928	350,000	15,881	74,556

Table 8.4.2: Breed cancer specific mortality for OS and mammary tumors. Listed are the mortality from tumors per breed, the rank among breeds for the development of OS and mammary tumors, the interquartile age for longevity, the AKC height, and breed standard weights. Note the trend for OS, the higher the height/weight the more likely the breeds are the develop OS.

Breed	Mortality from Tumors per 10,000 dog-years at risk [308]	Risk Rank Mammary Tumors [310]	Incidence Rate for Mammary Tumors [310]	Incident Rate Bone Tumors [311]	Breed Bone Tumor Rank [311]	Interquartile Range of Ages [280]	Size (kg) [312]	AKC Height (Inches) At Shoulders [313]
American cocker spaniel		5	192					13.5-15.5
Beagle		29	95				10	13-15
Bearded Collie		45	50			12.2-14.3		20-22
Belgian Tervuren		42	51				32	22-26
Bernese Mountain Dog	306	41	54	26	8		45	23-27.5
Bichon Frise		8	172			9.5–14.8		9.5-11.5
Border Collie		43	51			11.5-15	17	18-22
Border Terrier		24	116			8.9–13.1	6	11.5-15.5
Boxer	203	3	256	13	12	7–11.6	29	21.5-25
Cairn Terrier		3	148			10.6-15.5		10
Cavalier King Charles Spaniel	24	38	64			8.1–12.3	6	12-13

Dachshund (smooth/wired, normal size)	21	27	104			9.2-14.3 (Miniature)	7	
Dalmatian		36	72			11.5–14.0	25	19-23
Doberman Pinscher	168	2	297	24	10	6.2–11.0	35	24-28
Drever	29	39	62					
English Cocker Spaniel		13	142				14	15-17
English Springer Spaniel		1	319			10.4–14	23	19-20
Finnish Hound		50	17					
Finnish Spitz		47	40					15.5-20
Flat-Coated Retriever		31	94	35	7		29	22-24.5
German Pointer		16	128					23-25
German Shepherd dog	71	9	170			9.2–12.9	37	22-26
Golden Retriever	55	35	73	6	20	11.0–14.09	32	21.5-24
Great Dane	119			45	3	4.0–9.0	30	28-32
Greyhound	58			30	6	8.1–12.0		
Hovawart				28	9			
Irish setter		19	126	8	14			
Irish Wolfhound	296			99	1		54	
Jack Russell Terrier		26	105			9.3–15.7	7	
Jamthund		21	122					

Labrador Retriever	45	28	96			10.6–14.0	30	
Leonberger	197	7	176	53	4			
Lhasa Apso						7.7–15.3		
Miniature Dachshund	6	20	126					
Miniature Schnauzer		33	77					
Miniature/toy poodle		15	136			11.1-15.6		Miniature:>10 <15; Toy: <10
Mongrel	34	25	116					
Munsterlander		30	95					
Newfoundland	105	49	21	22	11		64	
Norwegian Elkhound		32	80					19.5-20.5
Nova Scotia Duck Tolling Retriever		17	126				20	17-21
Papillion		10	158					8-11
Petite basset Griffon		18	126					13-15
Pyrenees	108							25-32
Rottweiler		22	121	36	5	5.5–10.2	45	22-27
Rough-haired Collie		51	5			9.4–13.8		22-26
Saint Bernard	172			78	2			25.5.-27.5
Samoyed		37	72					19-23.5
Shetland Sheepdog		46	47			11.7–13.8		13-16
Shih-tzu		34	74			9.2–15.6		8-11

Soft Coated Wheaten Terrier		4	199					17-19
Springer Spaniel	44							19-20
Standard Poodle	18	23	120				25	>15
Standard Schnauzer		14	142	9	13			17.5-19.5
Swedish Hound		44	50					11.5-13.5
Tibetan Spaniel		48	29					10
Watchel dog		11	149					
Weimaraner						11.1–13.5	28	23-27
West Highland White Terrier		40	60			10.4-14.9		10-11
Yorkshire Terrier		6	188			10.0–15.1	3	

Table 8.4.3: Overrepresentation of specific cancers in specific breeds

Cancer	Breed	Odds Ratio (OR)	Relative Risk (RR)	Reference
Anal Sac Gland Carcinoma				
	Cavalier King Charles Spaniel	3.36		[314]
	English cocker spaniel	11		[314]
	English springer spaniel	3.65		[314]
	German Shepherd Dogs			[315], [316]
	Golden Retrievers			[315], [316]
Fibrosarcoma				
	Basset Hound		2.9	[317]
	Border Collie			[318]
	Doberman Pinscher		2.1	[317]
	English Setter		2.4	[317]
	German Shepherd		1.5	[317]
	Great Dane		2.6	[317]
	Golden Retriever	3.64		[318]
	Labrador Retriever		1.8	[317]

	Saint Bernard		3.7	[317]
Gastric Carcinoma				
	Belgian Shepherd			[319], [320]
	Bouvier des Flandres	36.5****		[321]
	Chow-Chow		46.2	[321], [322], [323]
	Collie/Rough Collie	26.1****		[319], [321]
	Groenendael (Belgian Shepherd Dog)	34.5****		[321]
	Norwegian Elkhound	6.1****		[321]
	Norwegian Lundehund			[324],[325]
	Staffordshire Bull Terrier			[319], [323]
	Standard poodle	7.6****		[321]
	Tervuren	56.1****		[324], [326]
Hemangiosarcoma				
	Border Collie			[318]
	Boxer	>4		[327], [328], [329], [330]
	German Shepherd	>3		[327], [331], [332]
	Golden Retrievers	6.15**		[333], [334]

		2.2***		[333], [335]
	Greyhound	>3		[327], [331]
	Irish Wolfhounds	13.1		[336]
	Italian Greyhounds	23.6		[336]
	Maltese dogs			[337]
	Miniature Dachshunds			[337]
	Vizslas	9.0*		[338]
	Whippets	13.7		[336]
Lymphoma				
	Airedale Terrier		2.7	[317]
	Basset Hounds		4.1	[317]
	Beagle		2.57	[339]
	Border Collie	1.32		[340]
	Bouvier des Flandres		2.31	[328], [339]
	Boxers	3.26		[340], [341]
	Bull Mastiff	6.76		[340], [342]
	Bulldog	4.73		[328], [340]

	Cocker Spaniels		1.15	[339], [343]
	Dobermans	2.75		[341], [340]
	English Springer Spaniel	1.27		[340]
	German Shepherds	2.03		[328], [341]
	Golden Retriever	2.35		[340], [344]
	Irish Setter	1.18		[340]
	Labrador retriever		1.7	[317]
	Miniature schnauzer	1.57		[340]
	Old English Sheepdog	2.24		[339], [340]
	Rottweiler	6.01		[340], [341]
	Saint Bernard		2.63	[339]
	Scottish Terrier		9.16	[328], [339]
	Staffordshire bullterrier	1.25		[340]
Malignant Histiocytosis/ Histiocytic sarcoma				
	Bernese Mountain Dog	45		[345], [346], [347], [348]
	Flat-Coated Retriever	62		[345], [349]
	Golden Retriever			[350]
	Labrador Retriever			[350]
	Pembroke Welsh Corgis	9.7		[351], [352]

	Rottweiler			[350]
Mast Cell Tumors				
	American Staffordshire	2.07		[328]
	Beagle			[353]
	Boston Terrier	4.21		[328], [354], [355]
	Boxer	10.24		[328], [354]
	Bullmastiff	3.6		[328]
	Chinese Shar-Pei	3.84		[328]
	Dutch Pug	3.41		[328]
	English Pointer	1.84		[328]
	English Setter	1.88		[328]
	Fox Terrier			[353], [355]
	Golden Retrievers	2.05		[328], [344]
	Labrador Retriever	2.22		[328]
	Rhodesian Ridgeback	5.07		[328]
	Staffordshire Bull Terrier	1.77		[355]
	Vizsla	4.84		[328], [338]
	Weimaraner	3.96		[328]
Melanoma				

	Airedale Terrier	3.41		[328]
	Beauce shepherd	2.38		[77]
	Boxer	4.35		[328]
	Chesapeake Bay Retriever	5.84		[328]
	Chinese Shar-Pei	24.43		[318]
	Chow-Chows	40.37		[318]
	Cocker Spaniel		1.65	[356]
	Doberman Pinscher	3.02		[328]
	German Shorthaired Pointer		2.5	[317]
	Golden Retriever	2.59		[328], [344]
	Gordon Setter		5.17	[356]
	Irish Setter	2.23		[328]
	Labrador retriever	1.72		[77]
	Miniature Schnauzer	7.53		[328]
	Poodle	1.63		[77]
	Rottweiler	1.17		[77]
	Schnauzer	1.93		[77]
	Scottish Terrier	3.07		[328], [77]
	Vizsla	17.34		[328]
Osteosarcoma				

	Beagle	3.8		[357]
	BMD			[311]
	Boxer			[311]
	Doberman Pinscher	2.3		[311], [17]
	Flat-Coated Retrievers			[345], [311]
	German Shepherd Dog	2.2		[357]
	Golden retriever	2.1		[311], [17]
	Great Dane	12		[311], [358]
	Greyhound	17.3		[311], [42]
	Hovawart			[311]
	Irish setter	3.5		[311], [17]
	Irish Wolfhound	20.7		[311], [17]
	Labrador Retriever	1.3		[17]
	Leonberger			[311], [359]
	Miniature Poodle	2.7		[357]
	Newfoundland			[311], [359]
	Rottweiler	14.6		[311], [358]
	Saint Bernard	11.9		[311], [17]
	Scottish Deerhound			[300]
	Standard schnauzer			[311]

Prostate Cancer				
	Airedale Terrier	2.46		[360]
	Beagle	1.49		[17], [360]
	Bernese Mountain Dog	2.41		[361]
	Bouveir des Flandres	5.51		[361]
	Doberman Pinscher	1.97		[360]
	German Shepherd Dog	2.6		[357]
	German Shorthaired Pointer	1.89		[360], [361]
	Norwegian Elkhound	3.28		[360]
	Scottish Terriers	3.81		[360], [361]
	Shetland Sheepdogs	1.82		[360]
Soft-Tissue Sarcoma				
	Boxer	1.77		[328]
	Golden Retrievers	2.88		[344], [328]
	Labrador Retriever	1.48		[328]
	Rhodesian Ridgeback	4.81		[328]
	Siberian Husky	2.68		[328]
Squamous cell Carcinoma				
	Basset Hound	3.97		[328]

	Dalmatian	6.94		[328]
	Labrador Retriever	2.41		[318]
	Poodles	4.61		[318]
	Samoyeds	24.63		[318]
	Scottish terrier		2.5	[317]
	Weimaraner		2.5	[317]
Squamous Cell Carcinoma of the Digit				
	Beauceron			[362]
	Black Standard Poodle	5.9		[363], [299]
	Briard	10.4		[362], [299]
	Dachshund			[364]
	Flat-Coated Retriever			[364]
	Giant Schnauzer	22.7		[299]
	Gordon Setter	11.1		[299]
	Kerry Blue Terrier	7.7		[299]
	Labrador Retrievers			[363]
	Rottweiler			[364]
Transitional Cell Carcinoma of the Bladder				

	Beagle	4.15		[365]
	Scottish terriers	18.9		[365]
	Shetland Sheepdogs	4.46		[366]
	West Highland White Terriers	3.02		[365]
	Wire-haired Fox Terriers	3.2		[365]

*Female gonadectomized all ages

**Scrotal

***Non-scrotal

****The PMR of gastric carcinoma by breed was calculated by dividing the number of gastric carcinomas in a breed by all tumors in the breed over the number of gastric carcinomas in all other breeds divided by all other tumors in the other breeds in the database [321].

Table 8.4.4 Comparative aspects of spontaneous cancers in humans and dogs

Osteosarcoma*

	Incidence or Prevalence Risk Factors	Histology	Biological Behavior	Treatment	Shared Molecular and Genetic Factors
Human	• 1000/year • Adolescent disease (peak onset 10-14 years)	• 85-95% high grade • Marked aneuploidy and karyotypic complexity	• 90% in the appendicular skeleton • Metaphyseal region of long bones (distal femur > proximal tibia > proximal humerus) • 85-90% clinically confined to primary site at presentation; 85-90% develop metastasis before 2 years without chemotherapy (lung > bone > soft tissues; regional lymph nodes <10%)	• Surgical amputation, limb-sparing procedures, or rotationplasty • Adjuvant multi-agent chemotherapy: doxorubicin, cisplatin, methotrexate • 5-year survival of 60-70% (nonmetastatic disease setting) or 10-30% (if metastasis found at initial diagnosis)	• *p53* gene mutations and loss of heterozygosity, p53 overexpression • *RB1* copy number loss, reduced RB1 protein expression • *PTEN* gene deletion, decreased PTEN expression • *MYC* copy number gain • IGF-1/IGF-1R activation enhances anchorage independent growth and invasion in OS cell lines • Constitutive activation of STAT3 associated with aggressive biological behavior • Abberant MET expression enhances migration in response to ligand (HGF); coexpression and heterodimerization of MET, EGFR, and Ron alters signal transduction and promotes resistance to targeted therapeutics • mTOR pathway activation enhances survival in OS cell lines • High ezrin expression associated with early metastasis and poor outcome
Canine	• >10,000/year • Middle-aged to older dogs (peak onset 7-9 years) • Increased inherited risk in Scottish Deerhounds, Rottweilers, greyhounds, Great Danes, Saint Bernards, Irish wolfhounds	• 95% high grade • Marked aneuploidy and karyotypic complexity	• 75% in the appendicular skeleton • Metaphyseal region of long bones (distal radius > proximal humerus > distal femur) • 85-90% clinically confined to primary site at presentation; 90% develop metastasis before 1 years without chemotherapy (lung > bone > soft tissues; regional lymph nodes 4.4%)	• Surgical amputation. limb-sparing procedures • Adjuvant therapy: Platinum-based (carboplatin, cisplatin) chemotherapy alone or in combination with doxorubicin • >50% of dogs do not live beyond 1 year postamputation; 90% die of disease by 2 years	

***References:** [7, 10, 28, 52, 58, 59, 65, 73, 74, 79, 80, 83, 149, 150, 163, 165, 190, 193, 194, 198, 262, 294, 296, 367]

Non-Hodgkins Lymphoma†

	Incidence or Prevalence Risk Factors	Histology	Biological Behavior	Treatment	Shared Molecular and Genetic Factors
Human	• 19.6/100,000	• Histologically classified according to the NCI Working Formulation • DLBCL (high grade) most common (30-44%) • Follicular (low grade) next most common (22.1%)	• Classified according to the WHO clinical staging system • Nodal or extranodal at presentation	• CHOP-like chemotherapy and rituximab (R-CHOP) • Involved field and/or total body irradiation • 5-year overall survival of 50-65% for DLBCL patients receiving R-CHOP chemotherapy	• Gene expression profiling supports the organization of canine and human DLBCL into molecular subtypes, activated B-cell (ABC) DLBCL and germinal center-like B-cell (GCB) DLBCL • NF-κB/p65 canonical pathway activation in DLBCL • Upregulation of miR-17-92 cluster • *PTEN* gene deletion (human GCB DLBCL, canine DLBCL cell lines) • *p53* gene mutations • *MYC* copy number gain and protein overexpression • *INK4A-ARF* gene deletion in human ABC DLBCL; deletion of CFA 11 (harboring *INK4A* locus) in canine high grade T-cell lymphoma • MYC-IgH translocations present in Burkitt's lymphoma
Canine	• 20-107/100,000 • Higher incidence in Boxers, bull mastiffs, basset hounds, St. Bernards, Scottish terriers, Golden Retrievers, Airedales, bulldogs	• Histologically classified according to the NCI Working Formulation • DLBCL (high grade) most common (36-58%) • 35-40% are T-cell malignancies	• Classified according to the modified WHO clinical staging system for domestic animals •Multicentric nodal presentation most common	• CHOP-based chemotherapy • Complete remission rates of 60-90% • Median survival 10-14 months for DLBCL; median survival 6-9 months for peripheral T-cell (high grade) malignancies	

†References: [7, 10, 28, 53, 65, 78, 82, 84, 85, 95, 98, 99, 101, 103, 104, 107, 109, 110, 112, 117, 121, 125, 126, 190-192, 213]

Bladder Cancer††

	Incidence or Prevalence Risk Factors	Histology	Biological Behavior	Treatment	Shared Molecular and Genetic Factors
Human	• >65,000/year • 2:1 ratio for the occurrence of bladder TCC in men versus women	• Graded according to the WHO/ISUP classification of human urothelial neoplasms • >65% low-grade noninfiltrative TCC • 20% high-grade invasive TCC	• Variable initial tumor location within the bladder • 50% of invasive TCC metastasize to regional lymph nodes, lungs, bone, other organs	• Low-grade TCC: transurethral resection and intravesical therapy can be curative • High-grade TCC: cystectomy, adjuvant chemotherapy (cisplatin, methotrexate, vinblastine, doxorubicin) • 5-year survival rates 78% (muscle-invasive lymph node-negative disease), 47% (extravesical lymph node-negative tumors) or 31% (lymph node-positive disease)	• Reduced expression of androgen receptor in higher-grade and –stage disease; high androgen receptor expression associated with less-aggressive forms of TCC • EGFR transcript and protein overexpression • p53 overexpression • Reduced or absent RB1 protein expression • Cyclooxygenase-2 overexpressed in invasive TCC and carcinoma *in situ* • High survivin expression and nuclear localization in TCC tumors • Telomerase activity detected in human TCC tumors and in urine samples from dogs with TCC • Overexpression of DNMT1 associated with more aggressive disease
Canine	• 2% of canine malignancies • 2:1 ratio for the occurrence of TCC in female versus male dogs • Higher incidence in Scottish terriers, West Highland white terriers, Shetland sheepdogs, beagles, wire hair fox terriers • Risk associated with obesity, insecticide and herbicide exposure	• Classified according to modified WHO/ISUP classification system for dogs • >90% infiltrative TCC • <10% non-infiltrative tumors	• Demonstrated risk associated with obesity, insecticide and herbicide exposure • Majority of canine TCC located in trigone region of the bladder with urethral (56%) and/or prostatic involvement (29%) • Metastasis present in 15-20% of dogs at diagnosis; 50% of dogs at death (regional lymph nodes; lung, bone, other organs, including the skin)	• Coarse fraction radiotherapy, intensity-modulated radiotherapy • Adjuvant chemotherapy: cyclooxygenase inhibitors, platinum agents (carboplatin, cisplatin), vinblastine • Median survival of 4-7 months with single-agent drug treatment; >8 months with multi-drug treatments	

††**References:** [7, 10, 28, 65, 67, 69, 72, 75, 213, 225, 365, 366, 368-370]

CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone; DLBCL, diffuse large B-cell lymphoma; DNMT1, DNA methyltransferase1; EGFR, epidermal growth factor receptor; HGF, hepatocyte growth factor; IGF-1, insulin-like growth factor-1; IGF-1R, IGF-1 receptor; INK4A-ARF, cyclin-dependent kinase inhibitor 2A; ISUP, International Society of Urologic Pathology; mTOR, mammalian target of rapamycin; NCI, National Cancer Institute; NFκB, nuclear factor kappa-B; PTEN, phosphatase and tensin homolog; RB1, retinoblastoma 1; STAT3, signal transducer and activator of transcription 3; TCC, transitional cell carcinoma; WHO, World Health Organization.

References

1. Bigham, A., et al., *Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data.* PLoS Genet, 2010. **6**(9): p. e1001116.
2. Varnado, C.L., et al., *Development of recombinant hemoglobin-based oxygen carriers.* Antioxid Redox Signal, 2013. **18**(17): p. 2314-28.
3. Bauer, D.E., S.C. Kamran, and S.H. Orkin, *Reawakening fetal hemoglobin: prospects for new therapies for the beta-globin disorders.* Blood, 2012. **120**(15): p. 2945-53.
4. Mazurier, C., L. Douay, and H. Lapillonne, *Red blood cells from induced pluripotent stem cells: hurdles and developments.* Curr Opin Hematol, 2011. **18**(4): p. 249-53.
5. Opazo, J.C., F.G. Hoffmann, and J.F. Storz, *Differential loss of embryonic globin genes during the radiation of placental mammals.* Proceedings of the National Academy of Sciences, 2008. **105**(35): p. 12950-12955.
6. Vukcevic, D., *Bayesian and Frequentist Methods and Analyses of Genome-Wide Association Studies*, in *Department of Statistics*. 2009, St. Hugh's College, University of Oxford. p. 240.
7. Alvarez, C.E., *Naturally occurring cancers in dogs: insights for translational genetics and medicine.* ILAR J, 2014. **55**(1): p. 16-45.
8. Khanna, C., et al., *The dog as a cancer model.* Nature biotechnology, 2006. **24**(9): p. 1065-1066.
9. Parker, H.G., A.L. Shearin, and E.A. Ostrander, *Man's best friend becomes biology's best in show: genome analyses in the domestic dog.* Annu Rev Genet, 2010. **44**: p. 309-36.
10. Rowell, J.L., D.O. McCarthy, and C.E. Alvarez, *Dog models of naturally occurring cancer.* Trends Mol Med, 2011. **17**(7): p. 380-8.
11. Churchill, G.A. and R.W. Doerge, *Empirical threshold values for quantitative trait mapping.* Genetics, 1994. **138**(3): p. 963-71.
12. Zaykin, D.V., S.S. Young, and P.H. Westfall, *Using the false discovery rate approach in the genetic dissection of complex traits: a response to Weller et al.* Genetics, 2000. **154**(4): p. 1917-8.
13. http://fbresearch.org/education/nobel-prize/.
14. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
15. Kim, K., et al., *Picking the most likely candidates for further development: Novel intersection-union tests for addressing multi-component hypotheses in comparative genomics.* Proceedings of the American Statistical Association, 2004. **ASA Section on ENAR Spring Meeting**: p. 1396-1402.
16. Mirabello, L., R.J. Troisi, and S.A. Savage, *Osteosarcoma incidence and survival rates from 1973 to 2004: data from the Surveillance, Epidemiology, and End Results Program.* Cancer, 2009. **115**(7): p. 1531-43.
17. Ru, G., B. Terracini, and L.T. Glickman, *Host related risk factors for canine osteosarcoma.* Vet J, 1998. **156**(1): p. 31-9.
18. Stephens, M. and D.J. Balding, *Bayesian statistical methods for genetic association studies.* Nat Rev Genet, 2009. **10**(10): p. 681-90.
19. Futreal, P.A., et al., *A census of human cancer genes.* Nat Rev Cancer, 2004. **4**(3): p. 177-83.
20. Davoli, T., et al., *Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome.* Cell, 2013. **155**(4): p. 948-62.

21. Zack, T.I., et al., *Pan-cancer patterns of somatic copy number alteration.* Nat Genet, 2013. **45**(10): p. 1134-1140.
22. Hallstrom, B.M. and A. Janke, *Mammalian evolution may not be strictly bifurcating.* Mol Biol Evol, 2010. **27**(12): p. 2804-16.
23. Lindblad-Toh, K., et al., *Genome sequence, comparative analysis and haplotype structure of the domestic dog.* Nature, 2005. **438**(7069): p. 803-19.
24. Rowell, J.L., D.O. McCarthy, and C.E. Alvarez, *Dog models of naturally occurring cancer.* Trends Mol Med, 2011.
25. Bromham, L., *The genome as a life-history character: why rate of molecular evolution varies between mammal species.* Philos Trans R Soc Lond B Biol Sci, 2011. **366**(1577): p. 2503-13.
26. Nicholas, F.W., A. Crook, and D.R. Sargan, *Internet resources cataloguing inherited disorders in dogs.* Vet J, 2011. **189**(2): p. 132-5.
27. Bellumori, T.P., et al., *Prevalence of inherited disorders among mixed-breed and purebred dogs: 27,254 cases (1995-2010).* J Am Vet Med Assoc, 2013. **242**(11): p. 1549-55.
28. Dobson, J.M., *Breed-predispositions to cancer in pedigree dogs.* ISRN Vet Sci, 2013. **2013**: p. 941275.
29. Freedman, A.H., et al., *Genome sequencing highlights the dynamic early history of dogs.* PLoS Genet, 2014. **10**(1): p. e1004016.
30. Gray, M.M., et al., *Linkage disequilibrium and demographic history of wild and domestic canids.* Genetics, 2009. **181**(4): p. 1493-505.
31. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
32. Lango Allen, H., et al., *Hundreds of variants clustered in genomic loci and biological pathways affect human height.* Nature, 2010. **467**(7317): p. 832-8.
33. Karlsson, E.K., et al., *Efficient mapping of mendelian traits in dogs through genome-wide association.* Nat Genet, 2007. **39**(11): p. 1321-8.
34. Rimbault, M., et al., *Derived variants at six genes explain nearly half of size reduction in dog breeds.* Genome Res, 2013. **23**(12): p. 1985-95.
35. Boyko, A.R., et al., *A simple genetic architecture underlies morphological variation in dogs.* PLoS Biol, 2010. **8**(8): p. e1000451.
36. Savage, S.A., et al., *Genome-wide association study identifies two susceptibility loci for osteosarcoma.* Nat Genet, 2013. **45**(7): p. 799-803.
37. Webster, M.T., R.S. Wells, and J.B. Clegg, *Analysis of variation in the human beta-globin gene cluster using a novel DHPLC technique.* Mutat Res, 2002. **501**(1-2): p. 99-103.
38. Alkan, C., B.P. Coe, and E.E. Eichler, *Genome structural variation discovery and genotyping.* Nat Rev Genet, 2011. **12**(5): p. 363-76.
39. Alvarez, C.E. and J.M. Akey, *Copy number variation in the domestic dog.* Mamm Genome, 2012. **23**(1-2): p. 144-63.
40. Chen, W.K., et al., *Mapping DNA structural variation in dogs.* Genome Res, 2009. **19**(3): p. 500-9.
41. Nicholas, T.J., et al., *The genomic architecture of segmental duplications and associated copy number variants in dogs.* Genome Res, 2009. **19**(3): p. 491-9.

42. Karlsson, E.K., et al., *Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B.* Genome Biol, 2013. **14**(12): p. R132.
43. Chari, R., et al., *Integrating the multiple dimensions of genomic and epigenomic landscapes of cancer.* Cancer Metastasis Rev, 2010. **29**(1): p. 73-93.
44. Ciriello, G., et al., *Emerging landscape of oncogenic signatures across human cancers.* Nat Genet, 2013. **45**(10): p. 1127-1133.
45. Roberts, S.A., et al., *An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers.* Nat Genet, 2013. **45**(9): p. 970-6.
46. Kumar, V., et al., *C57BL/6N mutation in Cytoplasmic FMRP interacting protein 2 regulates cocaine response.* Science, 2013. **342**(6165): p. 1508-12.
47. Lawrence, M.S., et al., *Mutational heterogeneity in cancer and the search for new cancer-associated genes.* Nature, 2013. **499**(7457): p. 214-8.
48. Kandoth, C., et al., *Mutational landscape and significance across 12 major cancer types.* Nature, 2013. **502**(7471): p. 333-9.
49. Cheon, D.J. and S. Orsulic, *Mouse models of cancer.* Annu Rev Pathol, 2011. **6**: p. 95-119.
50. Wei, W.H., G. Hemani, and C.S. Haley, *Detecting epistasis in human complex traits.* Nat Rev Genet, 2014.
51. Paoloni, M. and C. Khanna, *Translation of new cancer treatments from pet dogs to humans.* Nature reviews.Cancer, 2008. **8**(2): p. 147-156.
52. Scott, M.C., et al., *Molecular subtypes of osteosarcoma identified by reducing tumor heterogeneity through an interspecies comparative approach.* Bone, 2011. **49**(3): p. 356-67.
53. Mudaliar, M.A., et al., *Comparative gene expression profiling identifies common molecular signatures of NF-kappaB activation in canine and human diffuse large B cell lymphoma (DLBCL).* PLoS One, 2013. **8**(9): p. e72591.
54. Hogge, G.S., et al., *Development of human granulocyte-macrophage colony-stimulating factor-transfected tumor cell vaccines for the treatment of spontaneous canine cancer.* Human Gene Therapy, 1998. **9**(13): p. 1851-1861.
55. Ladiges, W.C., R. Storb, and E.D. Thomas, *Canine models of bone marrow transplantation.* Laboratory animal science, 1990. **40**(1): p. 11-15.
56. Kippenes, H., et al., *Spatial accuracy of fractionated IMRT delivery studies in canine paraspinal irradiation.* Veterinary radiology & ultrasound : the official journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association, 2003. **44**(3): p. 360-366.
57. Lawrence, J.A., et al., *Proof of principle of ocular sparing in dogs with sinonasal tumors treated with intensity-modulated radiation therapy.* Veterinary radiology & ultrasound : the official journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association, 2010. **51**(5): p. 561-570.
58. LaRue, S.M., et al., *Limb-sparing treatment for osteosarcoma in dogs.* Journal of the American Veterinary Medical Association, 1989. **195**(12): p. 1734-1744.
59. Withrow, S.J., et al., *Intra-arterial cisplatin with or without radiation in limb-sparing for canine osteosarcoma.* Cancer, 1993. **71**(8): p. 2484-2490.
60. Hershey, A.E., et al., *Inhalation chemotherapy for macroscopic primary or metastatic lung tumors: proof of principle using dogs with spontaneously occurring tumors as a*

model. Clinical cancer research : an official journal of the American Association for Cancer Research, 1999. **5**(9): p. 2653-2659.
61. American Veterinary Medical, A., *U.S. pet ownership & demographics sourcebook.* 2012.
62. https://ccrod.cancer.gov/confluence/pages/viewpage.action?pageId=46137977.
63. Adams, V.J., et al., *Methods and mortality results of a health survey of purebred dogs in the UK.* J Small Anim Pract, 2010. **51**(10): p. 512-24.
64. Bronson, R.T., *Variation in age at death of dogs of different sexes and breeds.* American Journal of Veterinary Research, 1982. **43**(11): p. 2057-2059.
65. Vail, D.M. and E.G. MacEwen, *Spontaneously occurring tumors of companion animals as models for human cancer.* Cancer investigation, 2000. **18**(8): p. 781-792.
66. Freiwald, A., A. Litster, and H.Y. Weng, *Survey to investigate pet ownership and attitudes to pet care in metropolitan Chicago dog and/or cat owners.* Prev Vet Med, 2014. **115**(3-4): p. 198-204.
67. Kelsey, J.L., A.S. Moore, and L.T. Glickman, *Epidemiologic studies of risk factors for cancer in pet dogs.* Epidemiologic reviews, 1998. **20**(2): p. 204-217.
68. Bukowski, J.A., D. Wartenberg, and M. Goldschmidt, *Environmental causes for sinonasal cancers in pet dogs, and their usefulness as sentinels of indoor cancer risk.* Journal of toxicology and environmental health.Part A, 1998. **54**(7): p. 579-591.
69. Hayes, H.M., Jr., R. Hoover, and R.E. Tarone, *Bladder cancer in pet dogs: a sentinel for environmental cancer?* American Journal of Epidemiology, 1981. **114**(2): p. 229-233.
70. Hayes, H.M., Jr. and J.F. Fraumeni, Jr., *Epidemiological features of canine renal neoplasms.* Cancer research, 1977. **37**(8 Pt 1): p. 2553-2556.
71. Hayes, H.H., Jr., *The comparative epidemiology of selected neoplasms between dogs, cats and humans. A review.* European journal of cancer, 1978. **14**(12): p. 1299-1308.
72. Glickman, L.T., et al., *Epidemiologic study of insecticide exposures, obesity, and risk of bladder cancer in household dogs.* Journal of toxicology and environmental health, 1989. **28**(4): p. 407-414.
73. Fenger, J.M., C.A. London, and W.C. Kisseberth, *Canine osteosarcoma: a naturally occurring disease to inform pediatric oncology.* ILAR J, 2014. **55**(1): p. 69-85.
74. Withrow, S.J. and R.M. Wilkins, *Cross talk from pets to people: translational osteosarcoma treatments.* ILAR journal / National Research Council, Institute of Laboratory Animal Resources, 2010. **51**(3): p. 208-213.
75. Knapp, D.W., et al., *Urinary bladder cancer in dogs, a naturally occurring model for cancer biology and drug development.* ILAR J, 2014. **55**(1): p. 100-18.
76. Liu, D., et al., *Molecular homology and difference between spontaneous canine mammary cancer and human breast cancer.* Cancer research, 2014.
77. Gillard, M., et al., *Naturally occurring melanomas in dogs as models for non-UV pathways of human melanomas.* Pigment Cell Melanoma Res, 2014. **27**(1): p. 90-102.
78. Marconato, L., M.E. Gelain, and S. Comazzi, *The dog as a possible animal model for human non-Hodgkin lymphoma: a review.* Hematol Oncol, 2013. **31**(1): p. 1-9.
79. Angstadt, A.Y., et al., *Characterization of canine osteosarcoma by array comparative genomic hybridization and RT-qPCR: signatures of genomic imbalance in canine osteosarcoma parallel the human counterpart.* Genes, chromosomes & cancer, 2011. **50**(11): p. 859-874.

80. Angstadt, A.Y., et al., *A genome-wide approach to comparative oncology: high-resolution oligonucleotide aCGH of canine and human osteosarcoma pinpoints shared microaberrations.* Cancer Genet, 2012. **205**(11): p. 572-87.
81. Paoloni, M., et al., *Canine tumor cross-species genomics uncovers targets linked to osteosarcoma progression.* BMC genomics, 2009. **10**: p. 625-2164-10-625.
82. Richards, K.L., et al., *Gene profiling of canine B-cell lymphoma reveals germinal center and postgerminal center subtypes with different survival times, modeling human DLBCL.* Cancer Res, 2013. **73**(16): p. 5029-39.
83. Fossey, S.L., et al., *Characterization of STAT3 activation and expression in canine and human osteosarcoma.* BMC cancer, 2009. **9**: p. 81-2407-9-81.
84. Breen, M. and J.F. Modiano, *Evolutionarily conserved cytogenetic changes in hematological malignancies of dogs and humans--man and his best friend share more than companionship.* Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology, 2008. **16**(1): p. 145-154.
85. Thomas, R., et al., *Refining tumor-associated aneuploidy through 'genomic recoding' of recurrent DNA copy number aberrations in 150 canine non-Hodgkin lymphomas.* Leukemia & lymphoma, 2011. **52**(7): p. 1321-1335.
86. London, C.A., et al., *Spontaneous canine mast cell tumors express tandem duplications in the proto-oncogene c-kit.* Experimental hematology, 1999. **27**(4): p. 689-697.
87. Letard, S., et al., *Gain-of-function mutations in the extracellular domain of KIT are common in canine mast cell tumors.* Molecular cancer research : MCR, 2008. **6**(7): p. 1137-1145.
88. Downing, S., et al., *Prevalence and importance of internal tandem duplications in exons 11 and 12 of c-kit in mast cell tumors of dogs.* American Journal of Veterinary Research, 2002. **63**(12): p. 1718-1723.
89. Gregory-Bryson, E., et al., *Canine and human gastrointestinal stromal tumors display similar mutations in c-KIT exon 11.* BMC cancer, 2010. **10**: p. 559-2407-10-559.
90. Liao, A.T., et al., *Inhibition of constitutively active forms of mutant kit by multitargeted indolinone tyrosine kinase inhibitors.* Blood, 2002. **100**(2): p. 585-593.
91. Pryer, N.K., et al., *Proof of target for SU11654: inhibition of KIT phosphorylation in canine mast cell tumors.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2003. **9**(15): p. 5729-5734.
92. London, C.A., et al., *Phase I dose-escalating study of SU11654, a small molecule receptor tyrosine kinase inhibitor, in dogs with spontaneous malignancies.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2003. **9**(7): p. 2755-2768.
93. Enginler, S.O., et al., *Genetic variations of BRCA1 and BRCA2 genes in dogs with mammary tumours.* Vet Res Commun, 2014. **38**(1): p. 21-7.
94. Rivera, P. and H. von Euler, *Molecular biological aspects on canine and human mammary tumors.* Vet Pathol, 2011. **48**(1): p. 132-46.
95. Ito, D., A.M. Frantz, and J.F. Modiano, *Canine lymphoma as a comparative model for human non-Hodgkin lymphoma: recent progress and applications.* Vet Immunol Immunopathol, 2014. **159**(3-4): p. 192-201.
96. http://www.cancer.gov/cancertopics/types/hodgkin.
97. http://www.cancer.gov/cancertopics/types/non-hodgkin.

98. Bosetti, C., et al., *Incidence and mortality from non-Hodgkin lymphoma in Europe: the end of an epidemic?* Int J Cancer, 2008. **123**(8): p. 1917-23.
99. Adamson, P., et al., *Time trends in the registration of Hodgkin and non-Hodgkin lymphomas in Europe.* Eur J Cancer, 2007. **43**(2): p. 391-401.
100. Costas, L., S. de Sanjose, and C. Infante-Rivard, *Reproductive factors and non-Hodgkin lymphoma: A systematic review.* Crit Rev Oncol Hematol, 2014.
101. Arico, A., et al., *The role of vascular endothelial growth factor and matrix metalloproteinases in canine lymphoma: in vivo and in vitro study.* BMC Vet Res, 2013. **9**: p. 94.
102. Sokolowska, J., et al., *Proliferation activity in canine lymphomas.* Pol J Vet Sci, 2012. **15**(4): p. 727-34.
103. Pinheiro, D., et al., *Dissecting the Regulatory Microenvironment of a Large Animal Model of Non-Hodgkin Lymphoma: Evidence of a Negative Prognostic Impact of FOXP3+ T Cells in Canine B Cell Lymphoma.* PLoS One, 2014. **9**(8): p. e105027.
104. Ito, D., et al., *A tumor-related lymphoid progenitor population supports hierarchical tumor organization in canine B-cell lymphoma.* J Vet Intern Med, 2011. **25**(4): p. 890-6.
105. Reimann-Berg, N., et al., *Cytogenetic analysis of CpG-oligonucleotide DSP30 plus Interleukin-2-Stimulated canine B-Cell lymphoma cells reveals the loss of one X Chromosome as the sole abnormality.* Cytogenet Genome Res, 2011. **135**(1): p. 79-82.
106. Frantz, A.M., et al., *Molecular profiling reveals prognostically significant subtypes of canine lymphoma.* Vet Pathol, 2013. **50**(4): p. 693-703.
107. Aresu, L., et al., *Canine indolent and aggressive lymphoma: clinical spectrum with histologic correlation.* Vet Comp Oncol, 2013.
108. Seifert, M., R. Scholtysik, and R. Kuppers, *Origin and pathogenesis of B cell lymphomas.* Methods Mol Biol, 2013. **971**: p. 1-25.
109. Martelli, M., et al., *Diffuse large B-cell lymphoma.* Crit Rev Oncol Hematol, 2013. **87**(2): p. 146-71.
110. Swerdlow, S.H.I.A.f.R.o.C.W.H.O., *WHO classification of tumours of haematopoietic and lymphoid tissues*. 2008, Lyon, France: International Agency for Research on Cancer.
111. Wang, E., et al., *Composite lymphoid neoplasm of B-cell and T-cell origins: a pathologic study of 14 cases.* Hum Pathol, 2014. **45**(4): p. 768-84.
112. Alizadeh, A.A., et al., *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.* Nature, 2000. **403**(6769): p. 503-11.
113. Lenz, G., et al., *Stromal gene signatures in large-B-cell lymphomas.* N Engl J Med, 2008. **359**(22): p. 2313-23.
114. Bohers, E., et al., *Activating somatic mutations in diffuse large B-cell lymphomas: lessons from next generation sequencing and key elements in the precision medicine era.* Leuk Lymphoma, 2014: p. 1-18.
115. Rosenwald, A., et al., *Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma.* J Exp Med, 2003. **198**(6): p. 851-62.
116. Oschlies, I., et al., *Diffuse large B-cell lymphoma in pediatric patients belongs predominantly to the germinal-center type B-cell lymphomas: a clinicopathologic analysis of cases included in the German BFM (Berlin-Frankfurt-Munster) Multicenter Trial.* Blood, 2006. **107**(10): p. 4047-52.

117. Lenz, G. and L.M. Staudt, *Aggressive lymphomas.* N Engl J Med, 2010. **362**(15): p. 1417-29.
118. Hatzi, K. and A. Melnick, *Breaking bad in the germinal center: how deregulation of BCL6 contributes to lymphomagenesis.* Trends Mol Med, 2014. **20**(6): p. 343-52.
119. Phan, R.T. and R. Dalla-Favera, *The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells.* Nature, 2004. **432**(7017): p. 635-9.
120. Phan, R.T., et al., *BCL6 interacts with the transcription factor Miz-1 to suppress the cyclin-dependent kinase inhibitor p21 and cell cycle arrest in germinal center B cells.* Nat Immunol, 2005. **6**(10): p. 1054-60.
121. Dunleavy, K. and W.H. Wilson, *Appropriate management of molecular subtypes of diffuse large B-cell lymphoma.* Oncology (Williston Park), 2014. **28**(4): p. 326-34.
122. Lossos, I.S., et al., *Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas.* Proc Natl Acad Sci U S A, 2000. **97**(18): p. 10209-13.
123. Ngo, V.N., et al., *Oncogenically active MYD88 mutations in human lymphoma.* Nature, 2011. **470**(7332): p. 115-9.
124. Ferraresso, S., et al., *Epigenetic silencing of TFPI-2 in canine diffuse large B-cell lymphoma.* PLoS One, 2014. **9**(4): p. e92707.
125. Habineza Ndikuyeze, G., et al., *A phase I clinical trial of systemically delivered NEMO binding domain peptide in dogs with spontaneous activated B-cell like diffuse large B-cell lymphoma.* PLoS One, 2014. **9**(5): p. e95404.
126. Sato, M., et al., *Evaluation of the prognostic significance of BCL6 gene expression in canine high-grade B-cell lymphoma.* Vet J, 2012. **191**(1): p. 108-14.
127. Ko, J.S., et al., *Primary lymphoma of the uterine horn in a Lhasa Apso dog.* Ir Vet J, 2013. **66**(1): p. 24.
128. Ueno, H., et al., *Extranodal lymphoma with peripheral nervous system involvement in a dog.* J Vet Med Sci, 2014. **76**(5): p. 723-7.
129. Nagel, C., T. Silver, and B. Grahn, *Optic chiasm B-cell lymphoma in a 20-month-old Mastiff dog.* Vet Ophthalmol, 2013. **16 Suppl 1**: p. 164-7.
130. Kol, A., et al., *B-cell lymphoma with plasmacytoid differentiation, atypical cytoplasmic inclusions, and secondary leukemia in a dog.* Vet Clin Pathol, 2013. **42**(1): p. 40-6.
131. Mueller, F., B. Fuchs, and B. Kaser-Hotz, *Comparative biology of human and canine osteosarcoma.* Anticancer Res, 2007. **27**(1A): p. 155-64.
132. Rankin, K.S., et al., *Of dogs and men: comparative biology as a tool for the discovery of novel biomarkers and drug development targets in osteosarcoma.* Pediatr Blood Cancer, 2012. **58**(3): p. 327-33.
133. http://www.ncin.org.uk/publications/data_briefings/bone_sarcomas_incidence_and_survival.
134. He, J.P., et al., *Review of the Molecular Pathogenesis of Osteosarcoma.* Asian Pac J Cancer Prev, 2014. **15**(15): p. 5967-5976.
135. Ottaviani, G. and N. Jaffe, *The etiology of osteosarcoma.* Cancer Treat Res, 2009. **152**: p. 15-32.
136. Ward, E.D.C.R.A.K.B.J.A., *Childhood and adolescent cancer statistics, 2014.* CAAC CA: A Cancer Journal for Clinicians, 2014. **64**(2): p. 83-103.

137. Kleinerman, E. *Current advances in osteosarcoma*. 2014; Available from: http://dx.doi.org/10.1007/978-3-319-04843-7.
138. Davis, B.W. and E.A. Ostrander, *Domestic dogs and cancer research: a breed-based genomics approach.* ILAR J, 2014. **55**(1): p. 59-68.
139. Wang, S., et al., *TFPI-2 is a putative tumor suppressor gene frequently inactivated by promoter hypermethylation in nasopharyngeal carcinoma.* BMC Cancer, 2010. **10**: p. 617.
140. Wong, C.M., et al., *Tissue factor pathway inhibitor-2 as a frequently silenced tumor suppressor gene in hepatocellular carcinoma.* Hepatology, 2007. **45**(5): p. 1129-38.
141. http://www.cancer.gov/cancertopics/types/bone.
142. Morello, E., M. Martano, and P. Buracco, *Biology, diagnosis and treatment of canine appendicular osteosarcoma: similarities and differences with human osteosarcoma.* Vet J, 2011. **189**(3): p. 268-77.
143. Selvarajah, G.T. and J. Kirpensteijn, *Prognostic and predictive biomarkers of canine osteosarcoma.* Vet J, 2010. **185**(1): p. 28-35.
144. Egenvall, A., et al., *Mortality in over 350,000 insured Swedish dogs from 1995-2000: II. Breed-specific age and survival patterns and relative risk for causes of death.* Acta Vet Scand, 2005. **46**(3): p. 121-36.
145. Longhi, A., et al., *Height as a risk factor for osteosarcoma.* J Pediatr Hematol Oncol, 2005. **27**(6): p. 314-8.
146. Boerman, I., et al., *Prognostic factors in canine appendicular osteosarcoma - a meta-analysis.* BMC Vet Res, 2012. **8**: p. 56.
147. Kalajzic, I., et al., *In vitro and in vivo approaches to study osteocyte biology.* Bone, 2013. **54**(2): p. 296-306.
148. Sottnik, J.L., et al., *Osteocytes serve as a progenitor cell of osteosarcoma.* J Cell Biochem, 2014. **115**(8): p. 1420-9.
149. Selvarajah, G.T., et al., *Gene expression profiling of canine osteosarcoma reveals genes associated with short and long survival times.* Mol Cancer, 2009. **8**: p. 72.
150. Rodriguez, C.O., Jr., *Using canine osteosarcoma as a model to assess efficacy of novel therapies: can old dogs teach us new tricks?* Adv Exp Med Biol, 2014. **804**: p. 237-56.
151. Shoeneman, J.K., et al., *Survivin inhibition via EZN-3042 in canine lymphoma and osteosarcoma.* Vet Comp Oncol, 2014.
152. http://www.ncbi.nlm.nih.gov/gene/332.
153. Tsai, H.C., et al., *CCN2 enhances resistance to cisplatin-mediating cell apoptosis in human osteosarcoma.* PLoS One, 2014. **9**(3): p. e90159.
154. Shoeneman, J.K., et al., *Expression and function of survivin in canine osteosarcoma.* Cancer Res, 2012. **72**(1): p. 249-59.
155. Hughes, D.P., *How the NOTCH pathway contributes to the ability of osteosarcoma cells to metastasize.* Cancer Treat Res, 2009. **152**: p. 479-96.
156. Zhang, P., et al., *Critical role of notch signaling in osteosarcoma invasion and metastasis.* Clin Cancer Res, 2008. **14**(10): p. 2962-9.
157. Dailey, D.D., et al., *HES1, a target of Notch signaling, is elevated in canine osteosarcoma, but reduced in the most aggressive tumors.* BMC Vet Res, 2013. **9**: p. 130.
158. Levine, R.A. and M.A. Fleischli, *Inactivation of p53 and retinoblastoma family pathways in canine osteosarcoma cell lines.* Vet Pathol, 2000. **37**(1): p. 54-61.

159. Thomas, R., et al., *Influence of genetic background on tumor karyotypes: evidence for breed-associated cytogenetic aberrations in canine appendicular osteosarcoma.* Chromosome Res, 2009. **17**(3): p. 365-77.
160. Levine, R.A., T. Forest, and C. Smith, *Tumor suppressor PTEN is mutated in canine osteosarcoma cell lines and tumors.* Vet Pathol, 2002. **39**(3): p. 372-8.
161. Flint, A.F., et al., *Overexpression of the erbB-2 proto-oncogene in canine osteosarcoma cell lines and tumors.* Vet Pathol, 2004. **41**(3): p. 291-6.
162. Liao, A.T., M. McMahon, and C. London, *Characterization, expression and function of c-Met in canine spontaneous cancers.* Vet Comp Oncol, 2005. **3**(2): p. 61-72.
163. Fossey, S.L., et al., *Oncostatin M promotes STAT3 activation, VEGF production, and invasion in osteosarcoma cell lines.* BMC Cancer, 2011. **11**: p. 125.
164. Gordon, I.K., F. Ye, and M.S. Kent, *Evaluation of the mammalian target of rapamycin pathway and the effect of rapamycin on target expression and cellular proliferation in osteosarcoma cells from dogs.* Am J Vet Res, 2008. **69**(8): p. 1079-84.
165. Hong, S.H., et al., *Protein kinase C regulates ezrin-radixin-moesin phosphorylation in canine osteosarcoma cells.* Vet Comp Oncol, 2011. **9**(3): p. 207-18.
166. Maniscalco, L., et al., *PDGFs and PDGFRs in canine osteosarcoma: new targets for innovative therapeutic strategies in comparative oncology.* Vet J, 2013. **195**(1): p. 41-7.
167. Loukopoulos, P., et al., *Characterisation of three novel canine osteosarcoma cell lines producing high levels of matrix metalloproteinases.* Res Vet Sci, 2004. **77**(2): p. 131-41.
168. Sarver, A.L., et al., *MicroRNAs at the human 14q32 locus have prognostic significance in osteosarcoma.* Orphanet J Rare Dis, 2013. **8**: p. 7.
169. Mortus, J.R., Y. Zhang, and D.P. Hughes, *Developmental pathways hijacked by osteosarcoma.* Adv Exp Med Biol, 2014. **804**: p. 93-118.
170. Fleming, J.M., K.E. Creevy, and D.E. Promislow, *Mortality in north american dogs from 1984 to 2004: an investigation into age-, size-, and breed-related causes of death.* J Vet Intern Med, 2011. **25**(2): p. 187-98.
171. Tjalma, R.A., *Canine bone sarcoma: estimation of relative risk as a function of body size.* J Natl Cancer Inst, 1966. **36**(6): p. 1137-50.
172. Ehrhart, N.P., S.D. Ryan, and T.M. Fan, *24 - Tumors of the Skeletal System*, in *Withrow and MacEwen's Small Animal Clinical Oncology (Fifth Edition)*, S.J.W.M.V.L. Page, Editor. 2013, W.B. Saunders: Saint Louis. p. 463-503.
173. Arora, R.S., et al., *Relationship between height at diagnosis and bone tumours in young people: a meta-analysis.* Cancer Causes Control, 2011. **22**(5): p. 681-8.
174. Mirabello, L., et al., *Height at diagnosis and birth-weight as risk factors for osteosarcoma.* Cancer Causes Control, 2011. **22**(6): p. 899-908.
175. Liu, J.P., et al., *Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r).* Cell, 1993. **75**(1): p. 59-72.
176. Hoopes, B.C., et al., *The insulin-like growth factor 1 receptor (IGF1R) contributes to reduced size in dogs.* Mamm Genome, 2012. **23**(11-12): p. 780-90.
177. van Duyvenvoorde, H.A., et al., *Copy number variants in patients with short stature.* Eur J Hum Genet, 2014. **22**(5): p. 602-9.
178. Fang, P., et al., *Severe short stature caused by novel compound heterozygous mutations of the insulin-like growth factor 1 receptor (IGF1R).* J Clin Endocrinol Metab, 2012. **97**(2): p. E243-7.

179. Klammt, J., W. Kiess, and R. Pfaffle, *IGF1R mutations as cause of SGA.* Best Pract Res Clin Endocrinol Metab, 2011. **25**(1): p. 191-206.
180. Maki, R.G., *Small is beautiful: insulin-like growth factors and their role in growth, development, and cancer.* J Clin Oncol, 2010. **28**(33): p. 4985-95.
181. Sutter, N.B., et al., *A single IGF1 allele is a major determinant of small size in dogs.* Science, 2007. **316**(5821): p. 112-5.
182. Kuijjer, M.L., et al., *IR/IGF1R signaling as potential target for treatment of high-grade osteosarcoma.* BMC Cancer, 2013. **13**: p. 245.
183. Maniscalco, L., et al., *Increased expression of insulin-like growth factor-1 receptor is correlated with worse survival in canine appendicular osteosarcoma.* Vet J, 2014.
184. Idelman, G., et al., *WT1-p53 interactions in insulin-like growth factor-I receptor gene regulation.* J Biol Chem, 2003. **278**(5): p. 3474-82.
185. Mirabello, L., et al., *A comprehensive candidate gene approach identifies genetic variation associated with osteosarcoma.* BMC Cancer, 2011. **11**: p. 209.
186. DiMasi, J.A., R.W. Hansen, and H.G. Grabowski, *The price of innovation: new estimates of drug development costs.* Journal of health economics, 2003. **22**(2): p. 151-185.
187. DiMasi, J.A. and H.G. Grabowski, *Economics of new oncology drug development.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2007. **25**(2): p. 209-216.
188. Kamb, A., S. Wee, and C. Lengauer, *Why is cancer drug discovery so difficult?* Nature reviews.Drug discovery, 2007. **6**(2): p. 115-120.
189. Kola, I. and J. Landis, *Can the pharmaceutical industry reduce attrition rates?* Nature reviews.Drug discovery, 2004. **3**(8): p. 711-715.
190. Hansen, K. and C. Khanna, *Spontaneous and genetically engineered animal models; use in preclinical cancer drug development.* European journal of cancer (Oxford, England : 1990), 2004. **40**(6): p. 858-880.
191. Valli, V.E., et al., *Classification of canine malignant lymphomas according to the World Health Organization criteria.* Veterinary pathology, 2011. **48**(1): p. 198-211.
192. Withrow, S.J.V.D.M.P.R.L., *Withrow & MacEwen's small animal clinical oncology*. 2013, Elsevier/Saunders: St. Louis, Mo.
193. Withrow, S.J. and C. Khanna, *Bridging the gap between experimental animals and humans in osteosarcoma.* Cancer treatment and research, 2009. **152**: p. 439-446.
194. Fan, T.M., *Animal models of osteosarcoma.* Expert review of anticancer therapy, 2010. **10**(8): p. 1327-1338.
195. Gordon, I., et al., *The Comparative Oncology Trials Consortium: using spontaneously occurring cancers in dogs to inform the cancer drug development pathway.* PLoS Med, 2009. **6**(10): p. e1000161.
196. Gordon, I.K. and C. Khanna, *Modeling opportunities in comparative oncology for drug development.* ILAR J, 2010. **51**(3): p. 214-20.
197. Khanna, C., et al., *Guiding the optimal translation of new cancer treatments from canine to human cancer patients.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2009. **15**(18): p. 5671-5677.
198. Paoloni, M.C., et al., *Rapamycin pharmacokinetic and pharmacodynamic relationships in osteosarcoma: a comparative oncology study in dogs.* PloS one, 2010. **5**(6): p. e11013.
199. Vail, D.M., *Cancer clinical trials: development and implementation.* The Veterinary clinics of North America.Small animal practice, 2007. **37**(6): p. 1033-57; v.

200. Burton, J. and C. Khanna, *The Role of Clinical Trials in Veterinary Oncology.* The Veterinary clinics of North America.Small animal practice, 2014. **44**(5): p. 977-987.
201. Khanna, C., et al., *Toward a drug development path that targets metastatic progression in osteosarcoma.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2014. **20**(16): p. 4200-4209.
202. Baneux, P.J., et al., *Issues related to institutional animal care and use committees and clinical trials using privately owned animals.* ILAR journal / National Research Council, Institute of Laboratory Animal Resources, 2014. **55**(1): p. 200-209.
203. McLemore, M.R., *The role of the data safety monitoring board: why was the Avastin phase III clinical trial stopped?* Clinical journal of oncology nursing, 2006. **10**(2): p. 153-154.
204. Zon, R., et al., *American Society of Clinical Oncology Statement on minimum standards and exemplary attributes of clinical trial sites.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2008. **26**(15): p. 2562-2567.
205. London, C.A., et al., *Phase I evaluation of STA-1474, a prodrug of the novel HSP90 inhibitor ganetespib, in dogs with spontaneous cancer.* PloS one, 2011. **6**(11): p. e27018.
206. Vail, D.M., et al., *Pegylated liposomal doxorubicin: proof of principle using preclinical animal models and pharmacokinetic studies.* Seminars in oncology, 2004. **31**(6 Suppl 13): p. 16-35.
207. Miller, A.B., et al., *Reporting results of cancer treatment.* Cancer, 1981. **47**(1): p. 207-214.
208. Eisenhauer, E.A., et al., *New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).* European journal of cancer (Oxford, England : 1990), 2009. **45**(2): p. 228-247.
209. Cheson, B.D., et al., *Revised response criteria for malignant lymphoma.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2007. **25**(5): p. 579-586.
210. Vail, D.M., et al., *Response evaluation criteria for peripheral nodal lymphoma in dogs (v1.0)--a Veterinary Cooperative Oncology Group (VCOG) consensus document.* Veterinary and comparative oncology, 2010. **8**(1): p. 28-37.
211. Nguyen, S.M., et al., *Response evaluation criteria for solid tumours in dogs (v1.0): a Veterinary Cooperative Oncology Group (VCOG) consensus document.* Veterinary and comparative oncology, 2013.
212. Lynch, S., et al., *Development of a questionnaire assessing health-related quality-of-life in dogs and cats with cancer.* Veterinary and comparative oncology, 2011. **9**(3): p. 172-182.
213. MacEwen, E.G., *Spontaneous tumors in dogs and cats: models for the study of cancer biology and treatment.* Cancer metastasis reviews, 1990. **9**(2): p. 125-136.
214. Rodriguez, C.O., Jr., et al., *Aerosol gemcitabine: preclinical safety and in vivo antitumor activity in osteosarcoma-bearing dogs.* Journal of aerosol medicine and pulmonary drug delivery, 2010. **23**(4): p. 197-206.
215. Khanna, C., et al., *Interleukin-2 liposome inhalation therapy is safe and effective for dogs with spontaneous pulmonary metastases.* Cancer, 1997. **79**(7): p. 1409-1421.
216. Khanna, C., et al., *Aerosol delivery of interleukin 2 liposomes is nontoxic and biologically effective: canine studies.* Clinical cancer research : an official journal of the American Association for Cancer Research, 1996. **2**(4): p. 721-734.

217. Khanna, C. and D.M. Vail, *Targeting the lung: preclinical and comparative evaluation of anticancer aerosols in dogs with naturally occurring cancers.* Current cancer drug targets, 2003. **3**(4): p. 265-273.
218. Stevenson, S., S.E. Emery, and V.M. Goldberg, *Factors affecting bone graft incorporation.* Clinical orthopaedics and related research, 1996. **(324)**(324): p. 66-74.
219. Stevenson, S., J.W. Shaffer, and V.M. Goldberg, *The humoral response to vascular and nonvascular allografts of bone.* Clinical orthopaedics and related research, 1996. **(326)**(326): p. 86-95.
220. Straw, R.C., et al., *The effect of intramedullary polymethylmethacrylate on healing of intercalary cortical allografts in a canine model.* Journal of orthopaedic research : official publication of the Orthopaedic Research Society, 1992. **10**(3): p. 434-439.
221. Mackie, T.R., et al., *Image guidance for precise conformal radiotherapy.* International journal of radiation oncology, biology, physics, 2003. **56**(1): p. 89-105.
222. Forrest, L.J., et al., *The utility of megavoltage computed tomography images from a helical tomotherapy system for setup verification purposes.* International journal of radiation oncology, biology, physics, 2004. **60**(5): p. 1639-1644.
223. Deveau, M.A., et al., *Dosimetric impact of daily setup variations during treatment of canine nasal tumors using intensity-modulated radiation therapy.* Veterinary radiology & ultrasound : the official journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association, 2010. **51**(1): p. 90-96.
224. Vaudaux, C., U. Schneider, and B. Kaser-Hotz, *Potential for intensity-modulated radiation therapy to permit dose escalation for canine nasal cancer.* Veterinary radiology & ultrasound : the official journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association, 2007. **48**(5): p. 475-481.
225. Nolan, M.W., et al., *Intensity-modulated and image-guided radiation therapy for treatment of genitourinary carcinomas in dogs.* Journal of veterinary internal medicine / American College of Veterinary Internal Medicine, 2012. **26**(4): p. 987-995.
226. Lawrence, J., et al., *Use of 3'-deoxy-3'-[18F]fluorothymidine PET/CT for evaluating response to cytotoxic chemotherapy in dogs with non-Hodgkin's lymphoma.* Veterinary radiology & ultrasound : the official journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association, 2009. **50**(6): p. 660-668.
227. LeBlanc, A.K., et al., *18FDG-PET imaging in canine lymphoma and cutaneous mast cell tumor.* Veterinary radiology & ultrasound : the official journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association, 2009. **50**(2): p. 215-223.
228. LeBlanc, A.K. and K. Peremans, *PET and SPECT imaging in veterinary medicine.* Seminars in nuclear medicine, 2014. **44**(1): p. 47-56.
229. Clausen, M.M., et al., *Dose escalation to high-risk sub-volumes based on non-invasive imaging of hypoxia and glycolytic activity in canine solid tumors: a feasibility study.* Radiation oncology (London, England), 2013. **8**: p. 262-717X-8-262.
230. Bowen, S.R., et al., *Spatially resolved regression analysis of pre-treatment FDG, FLT and Cu-ATSM PET from post-treatment FDG PET: an exploratory study.* Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology, 2012. **105**(1): p. 41-48.

231. Hansen, A.E., et al., *Multimodality functional imaging of spontaneous canine tumors using 64Cu-ATSM and 18FDG PET/CT and dynamic contrast enhanced perfusion CT.* Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology, 2012. **102**(3): p. 424-428.
232. Hansen, A.E., et al., *(64)Cu-ATSM and (18)FDG PET uptake and (64)Cu-ATSM autoradiography in spontaneous canine tumors: comparison with pimonidazole hypoxia immunohistochemistry.* Radiation oncology (London, England), 2012. **7**: p. 89-717X-7-89.
233. Bradshaw, T.J., et al., *Heterogeneity in intratumor correlations of 18F-FDG, 18F-FLT, and 61Cu-ATSM PET in canine sinonasal tumors.* Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 2013. **54**(11): p. 1931-1937.
234. Bradshaw, T.J., et al., *Spatiotemporal stability of Cu-ATSM and FLT positron emission tomography distributions during radiation therapy.* International journal of radiation oncology, biology, physics, 2014. **89**(2): p. 399-405.
235. Paoloni, M., et al., *Prospective molecular profiling of canine cancers provides a clinically relevant comparative model for evaluating personalized medicine (PMed) trials.* PloS one, 2014. **9**(3): p. e90028.
236. Kim, J.H., et al., *Interleukin-8 promotes canine hemangiosarcoma growth by regulating the tumor microenvironment.* Experimental cell research, 2014. **323**(1): p. 155-164.
237. Mitchell, L., et al., *Induction of remission results in spontaneous enhancement of anti-tumor cytotoxic T-lymphocyte activity in dogs with B cell lymphoma.* Veterinary immunology and immunopathology, 2012. **145**(3-4): p. 597-603.
238. Goulart, M.R., G.E. Pluhar, and J.R. Ohlfest, *Identification of myeloid derived suppressor cells in dogs with naturally occurring cancer.* PloS one, 2012. **7**(3): p. e33274.
239. Sherger, M., et al., *Identification of myeloid derived suppressor cells in the peripheral blood of tumor bearing dogs.* BMC veterinary research, 2012. **8**: p. 209-6148-8-209.
240. Mitchell, L., D.H. Thamm, and B.J. Biller, *Clinical and immunomodulatory effects of toceranib combined with low-dose cyclophosphamide in dogs with cancer.* Journal of veterinary internal medicine / American College of Veterinary Internal Medicine, 2012. **26**(2): p. 355-362.
241. Burton, J.H., et al., *Low-dose cyclophosphamide selectively decreases regulatory T cells and inhibits angiogenesis in dogs with soft tissue sarcoma.* Journal of veterinary internal medicine / American College of Veterinary Internal Medicine, 2011. **25**(4): p. 920-926.
242. Biller, B.J., et al., *Decreased ratio of CD8+ T cells to regulatory T cells associated with decreased survival in dogs with osteosarcoma.* Journal of veterinary internal medicine / American College of Veterinary Internal Medicine, 2010. **24**(5): p. 1118-1123.
243. Pardoll, D.M., *The blockade of immune checkpoints in cancer immunotherapy.* Nature reviews.Cancer, 2012. **12**(4): p. 252-264.
244. Lesterhuis, W.J., J.B. Haanen, and C.J. Punt, *Cancer immunotherapy--revisited.* Nature reviews.Drug discovery, 2011. **10**(8): p. 591-600.
245. Andersen, B.M., et al., *Vaccination for invasive canine meningioma induces in situ production of antibodies capable of antibody-dependent cell-mediated cytotoxicity.* Cancer research, 2013. **73**(10): p. 2987-2997.
246. Sorenmo, K.U., et al., *CD40-activated B cell cancer vaccine improves second clinical remission and survival in privately owned dogs with non-Hodgkin's lymphoma.* PloS one, 2011. **6**(8): p. e24167.

247. Riccardo, F., et al., *CSPG4-specific immunity and survival prolongation in dogs with oral malignant melanoma immunized with human CSPG4 DNA.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2014. **20**(14): p. 3753-3762.
248. Marconato, L., et al., *Randomized, placebo-controlled, double-blinded chemoimmunotherapy clinical trial in a pet dog model of diffuse large B-cell lymphoma.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2014. **20**(3): p. 668-677.
249. Gavazza, A., et al., *Safety and efficacy of a genetic vaccine targeting telomerase plus chemotherapy for the therapy of canine B-cell lymphoma.* Human Gene Therapy, 2013. **24**(8): p. 728-738.
250. Henson, M.S., et al., *Immunotherapy with autologous tumour antigen-coated microbeads (large multivalent immunogen), IL-2 and GM-CSF in dogs with spontaneous B-cell lymphoma.* Veterinary and comparative oncology, 2011. **9**(2): p. 95-105.
251. Singer, J., et al., *Comparative oncology: ErbB-1 and ErbB-2 homologues in canine cancer are susceptible to cetuximab and trastuzumab targeting.* Molecular immunology, 2012. **50**(4): p. 200-209.
252. Singer, J., et al., *Generation of a canine anti-EGFR (ErbB-1) antibody for passive immunotherapy in dog cancer patients.* Molecular cancer therapeutics, 2014. **13**(7): p. 1777-1790.
253. O'Connor, C.M., et al., *Adoptive T-cell therapy improves treatment of canine non-Hodgkin lymphoma post chemotherapy.* Scientific reports, 2012. **2**: p. 249.
254. Bergman, P.J., et al., *Long-term survival of dogs with advanced malignant melanoma after DNA vaccination with xenogeneic human tyrosinase: a phase I trial.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2003. **9**(4): p. 1284-1290.
255. Bergman, P.J., et al., *Development of a xenogeneic DNA vaccine program for canine malignant melanoma at the Animal Medical Center.* Vaccine, 2006. **24**(21): p. 4582-4585.
256. Liao, J.C., et al., *Vaccination with human tyrosinase DNA induces antibody responses in dogs with advanced melanoma.* Cancer immunity, 2006. **6**: p. 8.
257. MacEwen, E.G., et al., *Therapy for osteosarcoma in dogs with intravenous injection of liposome-encapsulated muramyl tripeptide.* Journal of the National Cancer Institute, 1989. **81**(12): p. 935-938.
258. Kurzman, I.D., et al., *Adjuvant therapy for osteosarcoma in dogs: results of randomized clinical trials using combined liposome-encapsulated muramyl tripeptide and cisplatin.* Clinical cancer research : an official journal of the American Association for Cancer Research, 1995. **1**(12): p. 1595-1601.
259. Meyers, P.A., et al., *Osteosarcoma: the addition of muramyl tripeptide to chemotherapy improves overall survival--a report from the Children's Oncology Group.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2008. **26**(4): p. 633-638.
260. Meyers, P.A. and A.J. Chou, *Muramyl tripeptide-phosphatidyl ethanolamine encapsulated in liposomes (L-MTP-PE) in the treatment of osteosarcoma.* Advances in Experimental Medicine and Biology, 2014. **804**: p. 307-321.

261. van Kempen, L.C. and W.P. Leenders, *Tumours can adapt to anti-angiogenic therapy depending on the stromal context: lessons from endothelial cell biology.* European journal of cell biology, 2006. **85**(2): p. 61-68.
262. Paoloni, M.C., et al., *Launching a novel preclinical infrastructure: comparative oncology trials consortium directed therapeutic targeting of TNFalpha to cancer vasculature.* PloS one, 2009. **4**(3): p. e4972.
263. Rusk, A., et al., *Preclinical evaluation of antiangiogenic thrombospondin-1 peptide mimetics, ABT-526 and ABT-510, in companion dogs with naturally occurring cancers.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2006. **12**(24): p. 7444-7455.
264. Rusk, A., et al., *Cooperative activity of cytotoxic chemotherapy with antiangiogenic thrombospondin-I peptides, ABT-526 in pet dogs with relapsed lymphoma.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2006. **12**(24): p. 7456-7464.
265. Sahora, A.I., et al., *Prospective study of thrombospondin-1 mimetic peptides, ABT-510 and ABT-898, in dogs with soft tissue sarcoma.* Journal of veterinary internal medicine / American College of Veterinary Internal Medicine, 2012. **26**(5): p. 1169-1176.
266. McCleese, J.K., et al., *The novel HSP90 inhibitor STA-1474 exhibits biologic activity against osteosarcoma cell lines.* International journal of cancer.Journal international du cancer, 2009. **125**(12): p. 2792-2801.
267. Lin, T.Y., et al., *The novel HSP90 inhibitor STA-9090 exhibits activity against Kit-dependent and -independent malignant mast cell tumors.* Experimental hematology, 2008. **36**(10): p. 1266-1277.
268. Moore, A.S., et al., *Doxorubicin and BAY 12-9566 for the treatment of osteosarcoma in dogs: a randomized, double-blind, placebo-controlled study.* Journal of veterinary internal medicine / American College of Veterinary Internal Medicine, 2007. **21**(4): p. 783-790.
269. Lequarre, A.S., et al., *LUPA: a European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs.* Vet J, 2011. **189**(2): p. 155-9.
270. http://nationalveterinarycancerregistry.org/.
271. http://texasvetcancerregistry.com/.
272. .vetcancersociety.org/, h.
273. .caninelifetimehealth.org/about/the-golden-retriever-lifetime-study/, h.
274. http://cdmrp.army.mil/pubs/annreports/2011annrep/2011annreport.pdf.
275. Hoeppner, M.P., et al., *An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts.* PloS one, 2014. **9**(3): p. e91172.
276. Vaysse, A., et al., *Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping.* PLoS Genet, 2011. **7**(10): p. e1002316.
277. Kellis, M., et al., *Defining functional DNA elements in the human genome.* Proc Natl Acad Sci U S A, 2014. **111**(17): p. 6131-8.
278. Guarente, L., *Aging research-where do we stand and where are we going?* Cell, 2014. **159**(1): p. 15-9.
279. Kraus, C., S. Pavard, and D.E. Promislow, *The size-life span trade-off decomposed: why large dogs die young.* Am Nat, 2013. **181**(4): p. 492-505.

280. O'Neill, D.G., et al., *Longevity and mortality of owned dogs in England.* Vet J, 2013. **198**(3): p. 638-43.
281. Hoffman, J.M., K.E. Creevy, and D.E. Promislow, *Reproductive capability is associated with lifespan and cause of death in companion dogs.* PLoS One, 2013. **8**(4): p. e61082.
282. Bonnett, B.N. and A. Egenvall, *Age patterns of disease and death in insured Swedish dogs, cats and horses.* J Comp Pathol, 2010. **142 Suppl 1**: p. S33-8.
283. Briggs, J., et al., *A compendium of canine normal tissue gene expression.* PloS one, 2011. **6**(5): p. e17107.
284. Rissetto, K.C., et al., *Cloning and expression of canine CD25 for validation of an anti-human CD25 antibody to compare T regulatory lymphocytes in healthy dogs and dogs with osteosarcoma.* Veterinary immunology and immunopathology, 2010. **135**(1-2): p. 137-145.
285. Biller, B.J., et al., *Use of FoxP3 expression to identify regulatory T cells in healthy dogs and dogs with cancer.* Veterinary immunology and immunopathology, 2007. **116**(1-2): p. 69-78.
286. Biller, B., *Metronomic Chemotherapy in Veterinary Patients with Cancer: Rethinking the Targets and Strategies of Chemotherapy.* The Veterinary clinics of North America.Small animal practice, 2014. **44**(5): p. 817-829.
287. LeBlanc, A.K., *Cancer and comparative imaging.* ILAR journal / National Research Council, Institute of Laboratory Animal Resources, 2014. **55**(1): p. 164-168.
288. Leblanc, A.K., et al., *Preliminary evaluation of serial (18) FDG-PET/CT to assess response to toceranib phosphate therapy in canine cancer.* Veterinary radiology & ultrasound : the official journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association, 2012. **53**(3): p. 348-357.
289. Vail, D.M., et al., *Assessment of GS-9219 in a pet dog model of non-Hodgkin's lymphoma.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2009. **15**(10): p. 3503-3510.
290. Statham-Ringen, K.A., et al., *Evaluation of a B-cell leukemia-lymphoma 2-specific radiolabeled peptide nucleic acid-peptide conjugate for scintigraphic detection of neoplastic lymphocytes in dogs with B-cell lymphoma.* American Journal of Veterinary Research, 2012. **73**(5): p. 681-688.
291. Eward, W.C., et al., *A novel imaging system permits real-time in vivo tumor bed assessment after resection of naturally occurring sarcomas in dogs.* Clinical orthopaedics and related research, 2013. **471**(3): p. 834-842.
292. Madajewski, B., et al., *Intraoperative near-infrared imaging of surgical wounds after tumor resections can detect residual disease.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2012. **18**(20): p. 5741-5751.
293. MacEwen, E.G., et al., *Adjuvant therapy for melanoma in dogs: results of randomized clinical trials using surgery, liposome-encapsulated muramyl tripeptide, and granulocyte macrophage colony-stimulating factor.* Clinical cancer research : an official journal of the American Association for Cancer Research, 1999. **5**(12): p. 4249-4258.
294. Khanna, C., et al., *A randomized controlled trial of octreotide pamoate long-acting release and carboplatin versus carboplatin alone in dogs with naturally occurring osteosarcoma: evaluation of insulin-like growth factor suppression and chemotherapy.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2002. **8**(7): p. 2406-2412.

295. Couto, J.I., et al., *Biologic activity of the novel small molecule STAT3 inhibitor LLL12 against canine osteosarcoma cell lines.* BMC veterinary research, 2012. **8**: p. 244-6148-8-244.
296. Khanna, C., et al., *The membrane-cytoskeleton linker ezrin is necessary for osteosarcoma metastasis.* Nature medicine, 2004. **10**(2): p. 182-186.
297. Onimoe, G.I., et al., *Small molecules, LLL12 and FLLL32, inhibit STAT3 and exhibit potent growth suppressive activity in osteosarcoma cells and tumor growth in mice.* Investigational new drugs, 2012. **30**(3): p. 916-926.
298. Jones, P., et al., *Single-nucleotide-polymorphism-based association mapping of dog stereotypes.* Genetics, 2008. **179**(2): p. 1033-44.
299. Karyadi, D.M., et al., *A copy number variant at the KITLG locus likely confers risk for canine squamous cell carcinoma of the digit.* PLoS Genet, 2013. **9**(3): p. e1003409.
300. Phillips, J.C., et al., *Heritability and segregation analysis of osteosarcoma in the Scottish deerhound.* Genomics, 2007. **90**(3): p. 354-63.
301. Shearin, A.L., et al., *The MTAP-CDKN2A locus confers susceptibility to a naturally occurring canine cancer.* Cancer Epidemiol Biomarkers Prev, 2012. **21**(7): p. 1019-27.
302. Szreter, S. and M. Woolcock, *Health by association? Social capital, social theory, and the political economy of public health.* Int J Epidemiol, 2004. **33**(4): p. 650-67.
303. Ransohoff, D.F., *Bias as a threat to the validity of cancer molecular-marker research.* Nat Rev Cancer, 2005. **5**(2): p. 142-9.
304. Yabroff, K.R., et al., *Economic burden of cancer in the United States: estimates, projections, and future research.* Cancer Epidemiol Biomarkers Prev, 2011. **20**(10): p. 2006-14.
305. Klein, U. and R. Dalla-Favera, *Germinal centres: role in B-cell physiology and malignancy.* Nat Rev Immunol, 2008. **8**(1): p. 22-33.
306. Basso, K. and R. Dalla-Favera, *Roles of BCL6 in normal and transformed germinal center B cells.* Immunol Rev, 2012. **247**(1): p. 172-83.
307. Bonnett, B.N., et al., *Mortality in insured Swedish dogs: rates and causes of death in various breeds.* Vet Rec, 1997. **141**(2): p. 40-4.
308. Bonnett, B.N., et al., *Mortality in over 350,000 insured Swedish dogs from 1995-2000: I. Breed-, gender-, age- and cause-specific rates.* Acta Vet Scand, 2005. **46**(3): p. 105-20.
309. Proschowsky, H.F., H. Rugbjerg, and A.K. Ersboll, *Mortality of purebred and mixed-breed dogs in Denmark.* Prev Vet Med, 2003. **58**(1-2): p. 63-74.
310. Egenvall, A., et al., *Incidence of and survival after mammary tumors in a population of over 80,000 insured female dogs in Sweden from 1995 to 2002.* Prev Vet Med, 2005. **69**(1-2): p. 109-27.
311. Egenvall, A., A. Nodtvedt, and H. von Euler, *Bone tumors in a population of 400 000 insured Swedish dogs up to 10 y of age: incidence and survival.* Can J Vet Res, 2007. **71**(4): p. 292-9.
312. Vaysse, A., et al., *Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping.* PLoS Genet, 2011. **7**(10): p. e1002316.
313. https://www.akc.org/.
314. Polton, G.A., et al., *Breed, gender and neutering status of British dogs with anal sac gland carcinoma.* Vet Comp Oncol, 2006. **4**(3): p. 125-31.
315. Williams, L.E., et al., *Carcinoma of the apocrine glands of the anal sac in dogs: 113 cases (1985-1995).* J Am Vet Med Assoc, 2003. **223**(6): p. 825-31.

316. Polton, G.A. and M.J. Brearley, *Clinical stage, therapy, and prognosis in canine anal sac gland carcinoma.* J Vet Intern Med, 2007. **21**(2): p. 274-80.
317. Priester, W.A., et al., *The occurrence of tumors in domestic animals*. National Cancer Institute monograph;54. 1980, Bethesda, Md. : Washington, D.C.: U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute ; For sale only by the Supt. of Docs., U.S. G.P.O. xi, 210 p.
318. Dennis, M.M., et al., *Frequency of and risk factors associated with lingual lesions in dogs: 1,196 cases (1995-2004).* J Am Vet Med Assoc, 2006. **228**(10): p. 1533-7.
319. Gualtieri, M., M.G. Monzeglio, and E. Scanziani, *Gastric neoplasia.* Vet Clin North Am Small Anim Pract, 1999. **29**(2): p. 415-40.
320. Sautter, J.H. and G.F. Hanlon, *Gastric neoplasms in the dog: a report of 20 cases.* J Am Vet Med Assoc, 1975. **166**(7): p. 691-6.
321. Seim-Wikse, T., et al., *Breed predisposition to canine gastric carcinoma--a study based on the Norwegian canine cancer register.* Acta Vet Scand, 2013. **55**: p. 25.
322. Penninck, D.G., A.S. Moore, and J. Gliatto, *Ultrasonography of canine gastric epithelial neoplasia.* Vet Radiol Ultrasound, 1998. **39**(4): p. 342-8.
323. Sullivan, M., et al., *A study of 31 cases of gastric carcinoma in dogs.* Vet Rec, 1987. **120**(4): p. 79-83.
324. Seim-Wikse, T., et al., *Tumour gastrin expression and serum gastrin concentrations in dogs with gastric carcinoma are poor diagnostic indicators.* J Comp Pathol, 2014. **151**(2-3): p. 207-11.
325. Meuten, D.J., *Tumors in domestic animals*. 2002, Ames, Iowa: Iowa State Press.
326. Lubbes, D., et al., *[Incidence of gastric carcinoma in Dutch Tervueren shepherd dogs born between 1991 and 2002]*. Tijdschr Diergeneeskd, 2009. **134**(14-15): p. 606-10.
327. Tamburini, B.A., et al., *Gene expression profiles of sporadic canine hemangiosarcoma are uniquely associated with breed.* PLoS One, 2009. **4**(5): p. e5549.
328. Villamil, J.A., et al., *Identification of the most common cutaneous neoplasms in dogs and evaluation of breed and age distributions for selected neoplasms.* J Am Vet Med Assoc, 2011. **239**(7): p. 960-5.
329. Sargan, D., *Haemangiosarcoma in German shepherd dogs.* Vet Rec, 2007. **160**(19): p. 672.
330. Ng, C.Y. and J.N. Mills, *Clinical and haematological features of haemangiosarcoma in dogs.* Aust Vet J, 1985. **62**(1): p. 1-4.
331. Srebernik, N. and E.C. Appleby, *Breed prevalence and sites of haemangioma and haemangiosarcoma in dogs.* Vet Rec, 1991. **129**(18): p. 408-9.
332. Cole, P.A., *Association of canine splenic hemangiosarcomas and hematomas with nodular lymphoid hyperplasia or siderotic nodules.* J Vet Diagn Invest, 2012. **24**(4): p. 759-62.
333. Trappler, M.C., et al., *Scrotal tumors in dogs: a retrospective study of 676 cases (1986-2010).* Can Vet J, 2014. **55**(1): p. 1229-33.
334. Prymak, C., et al., *Epidemiologic, clinical, pathologic, and prognostic characteristics of splenic hemangiosarcoma and splenic hematoma in dogs: 217 cases (1985).* J Am Vet Med Assoc, 1988. **193**(6): p. 706-12.
335. http://idid.vet.cam.ac.uk/detail.php?record=1563.
336. Goldschmidt, M.H.S.F.S., *Skin tumors of the dog and cat.* 1992, Oxford [England]; New York: Pergamon Press.

337. Yamamoto, S., et al., *Epidemiological, clinical and pathological features of primary cardiac hemangiosarcoma in dogs: a review of 51 cases.* J Vet Med Sci, 2013. **75**(11): p. 1433-41.
338. Zink, M.C., et al., *Evaluation of the risk and age of onset of cancer and behavioral disorders in gonadectomized Vizslas.* J Am Vet Med Assoc, 2014. **244**(3): p. 309-19.
339. Teske, E., *Canine malignant lymphoma: a review and comparison with human non-Hodgkin's lymphoma.* Vet Q, 1994. **16**(4): p. 209-19.
340. Edwards, D.S., et al., *Breed incidence of lymphoma in a UK population of insured dogs.* Vet Comp Oncol, 2003. **1**(4): p. 200-6.
341. Jagielski, D., et al., *A retrospective study of the incidence and prognostic factors of multicentric lymphoma in dogs (1998-2000).* J Vet Med A Physiol Pathol Clin Med, 2002. **49**(8): p. 419-24.
342. Onions, D.E., *A prospective survey of familial canine lymphosarcoma.* J Natl Cancer Inst, 1984. **72**(4): p. 909-12.
343. Pastor, M., et al., *Genetic and environmental risk indicators in canine non-Hodgkin's lymphomas: breed associations and geographic distribution of 608 cases diagnosed throughout France over 1 year.* J Vet Intern Med, 2009. **23**(2): p. 301-10.
344. Boerkamp, K.M., et al., *Estimated incidence rate and distribution of tumours in 4,653 cases of archival submissions derived from the Dutch golden retriever population.* BMC Vet Res, 2014. **10**: p. 34.
345. Erich, S.A., G.R. Rutteman, and E. Teske, *Causes of death and the impact of histiocytic sarcoma on the life expectancy of the Dutch population of Bernese mountain dogs and Flat-coated retrievers.* Vet J, 2013. **198**(3): p. 678-83.
346. Nielsen, L., et al., *Malignant histiocytosis and other causes of death in Bernese mountain dogs in Denmark.* Vet Rec, 2010. **166**(7): p. 199-202.
347. Constantino-Casas, F., et al., *The clinical presentation and histopathologic-immunohistochemical classification of histiocytic sarcomas in the Flat Coated Retriever.* Vet Pathol, 2011. **48**(3): p. 764-71.
348. Dobson, J., et al., *Mortality in a cohort of flat-coated retrievers in the UK.* Vet Comp Oncol, 2009. **7**(2): p. 115-21.
349. Morris, J.S., et al., *Histopathological survey of neoplasms in flat-coated retrievers, 1990 to 1998.* Vet Rec, 2000. **147**(11): p. 291-5.
350. Affolter, V.K. and P.F. Moore, *Localized and disseminated histiocytic sarcoma of dendritic cell origin in dogs.* Vet Pathol, 2002. **39**(1): p. 74-83.
351. Ide, T., et al., *Pathological and immunohistochemical features of subdural histiocytic sarcomas in 15 dogs.* J Vet Diagn Invest, 2011. **23**(1): p. 127-32.
352. Takahashi, M., et al., *Clinical characteristics and prognostic factors in dogs with histiocytic sarcomas in Japan.* J Vet Med Sci, 2014. **76**(5): p. 661-6.
353. Rothwell, T.L., et al., *Skin neoplasms of dogs in Sydney.* Aust Vet J, 1987. **64**(6): p. 161-4.
354. Brodey, R.S., *Canine and feline neoplasia.* Adv Vet Sci Comp Med, 1970. **14**: p. 309-54.
355. Warland, J. and J. Dobson, *Breed predispositions in canine mast cell tumour: a single centre experience in the United Kingdom.* Vet J, 2013. **197**(2): p. 496-8.
356. Ramos-Vara, J.A., et al., *Retrospective study of 338 canine oral melanomas with clinical, histologic, and immunohistochemical review of 129 cases.* Vet Pathol, 2000. **37**(6): p. 597-608.

357. Langenbach, A., et al., *Extraskeletal osteosarcomas in dogs: a retrospective study of 169 cases (1986-1996).* J Am Anim Hosp Assoc, 1998. **34**(2): p. 113-20.
358. Rosenberger, J.A., N.V. Pablo, and P.C. Crawford, *Prevalence of and intrinsic risk factors for appendicular osteosarcoma in dogs: 179 cases (1996-2005).* J Am Vet Med Assoc, 2007. **231**(7): p. 1076-80.
359. Anfinsen, K.P., et al., *Breed-specific incidence rates of canine primary bone tumors--a population based survey of dogs in Norway.* Can J Vet Res, 2011. **75**(3): p. 209-15.
360. Bryan, J.N., et al., *A population study of neutering status as a risk factor for canine prostate cancer.* Prostate, 2007. **67**(11): p. 1174-81.
361. Teske, E., et al., *Canine prostate carcinoma: epidemiological evidence of an increased risk in castrated dogs.* Mol Cell Endocrinol, 2002. **197**(1-2): p. 251-5.
362. Belluco, S., et al., *Digital squamous cell carcinoma in dogs: epidemiological, histological, and immunohistochemical study.* Vet Pathol, 2013. **50**(6): p. 1078-82.
363. O'Brien, M.G., J. Berg, and S.J. Engler, *Treatment by digital amputation of subungual squamous cell carcinoma in dogs: 21 cases (1987-1988).* J Am Vet Med Assoc, 1992. **201**(5): p. 759-61.
364. Wobeser, B.K., et al., *Diagnoses and clinical outcomes associated with surgically amputated canine digits submitted to multiple veterinary diagnostic laboratories.* Vet Pathol, 2007. **44**(3): p. 355-61.
365. Knapp, D.W. and S.K. McMillan, *29 - Tumors of the Urinary System*, in *Withrow and MacEwen's Small Animal Clinical Oncology (Fifth Edition)*, S.J.W.M.V.L. Page, Editor. 2013, W.B. Saunders: Saint Louis. p. 572-582.
366. Hanazono, K., et al., *Epidermal Growth Factor Receptor Expression in Canine Transitional Cell Carcinoma.* J Vet Med Sci, 2014.
367. Withrow, S.J.V.D.M.P.R.L. *Withrow & MacEwen's small animal clinical oncology*. 2013; Available from: http://www.sciencedirect.com/science/book/9781437723625.
368. Berent, A.C., et al., *Use of indwelling, double-pigtail stents for treatment of malignant ureteral obstruction in dogs: 12 cases (2006-2009).* Journal of the American Veterinary Medical Association, 2011. **238**(8): p. 1017-1025.
369. Blackburn, A.L., et al., *Evaluation of outcome following urethral stent placement for the treatment of obstructive carcinoma of the urethra in dogs: 42 cases (2004-2008).* Journal of the American Veterinary Medical Association, 2013. **242**(1): p. 59-68.
370. Weisse, C., et al., *Evaluation of palliative stenting for management of malignant urethral obstructions in dogs.* Journal of the American Veterinary Medical Association, 2006. **229**(2): p. 226-234.

Appendix 11

Presentation: OSU Humanities & Cognitive Sciences High School Summer Institute

OSU Humanities & Cognitive Sciences High School Summer Institute, lecture on behavioral and canine genetics (Aug 13, 2015)

Why we do what we do?: Disentangling the Threads of Programmed and Learned Behavior

Carlos E. Alvarez, PhD

Humanities & Cognitive Sciences
High School Summer Institute
August 13, 2015

Nationwide Children's Hospital
The Research Institute
The Ohio State University Medical Center

1. Meaning of life

2. Evolution of emotions

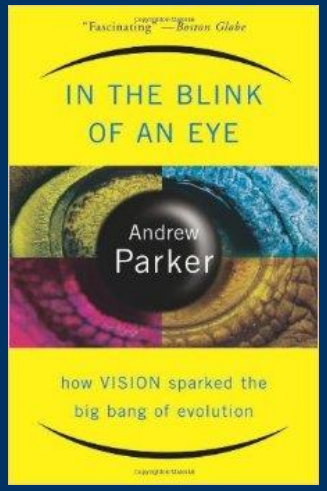
3. Why & how to study emotions

1. What is the meaning of life?

2. Evolution of emotions: origins

If procreation is the meaning of life,
our goal is survival until having babies, ∴

- ✓ **Eat**
- ✓ **Don't be killed**
- ✓ ***Mate!***

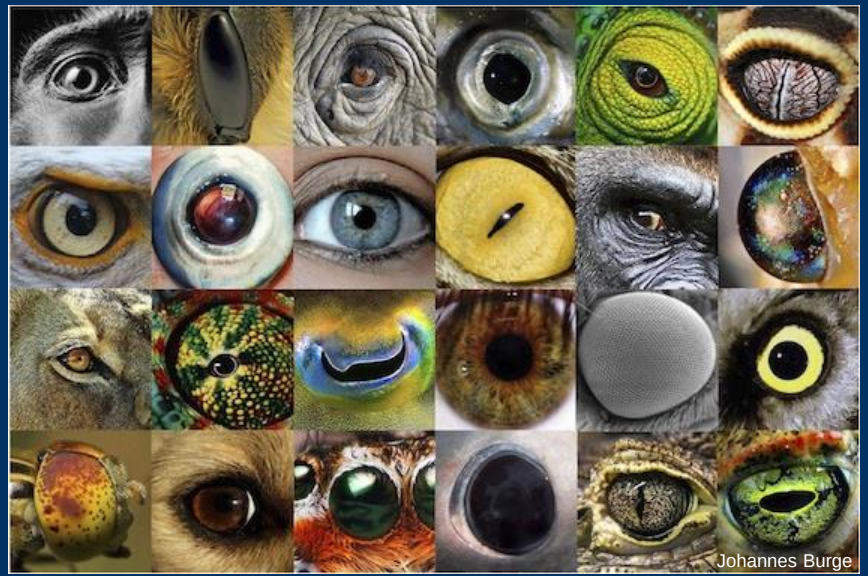


545 MYA: 38 types of organisms (phyla)

550 MYA: 3 phyla


PORIFERANS
CTENOPHORES
CNIDARIANS
ACOELOMATES
PSEUDOCOELOMATES
MOLLUSCS
ARTHROPODS
ONYCHOPHORANS
ANNELIDS
TARDIGRADES
BRACHIOPODS
PRIAPULIDS
HEMICHORDATES
CHORDATES
ECHINODERMS
POGONOPHORANS
CHAETOGNATHES
Cambrian Radiation
0
CENOZOIC
-65
MESOZOIC
-251
PALAEOZOIC
-542 Ma
NEOPROTEROZO

How important is vision?



- 6 of 38 phyla on earth have eyes
- Those 6 represent 95% of all living animals

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 12278–12282, October 1996
Biochemistry

Novel $G_q\alpha$ isoform is a candidate transducer of rhodopsin signaling in a *Drosophila* testes-autonomous pacemaker

(G protein/phototransduction/alternative splicing)

CARLOS E. ALVAREZ, KEITH ROBISON, AND WALTER GILBERT*

Department of Molecular and Cellular Biology, The Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138

BMC Evolutionary Biology

BioMed Central

Research article

Open Access

On the origins of arrestin and rhodopsin

Carlos E Alvarez[1,2,3]

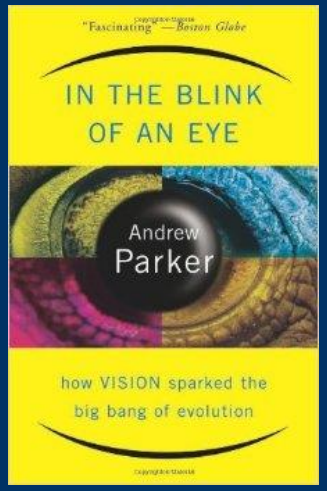
Address: [1]Center for Molecular and Human Genetics, The Research Institute at Nationwide Children's Hospital, Columbus, OH, 43205, USA, [2]Department of Pediatrics, The Ohio State University College of Medicine, Columbus, OH, 43210, USA and [3]Novartis Institutes of BioMedical Research, CH-4002 Basel, Switzerland

Email: Carlos E Alvarez - carlos.alvarez@nationwidechildrens.org

Published: 29 July 2008

BMC Evolutionary Biology 2008, **8**:222 doi:10.1186/1471-2148-8-222

Received: 11 January 2008
Accepted: 29 July 2008



545 MYA: 38 types of organisms (phyla)

550 MYA: 3 phyla

3. Why and how to study emotion: *what is it?*

Emotion:

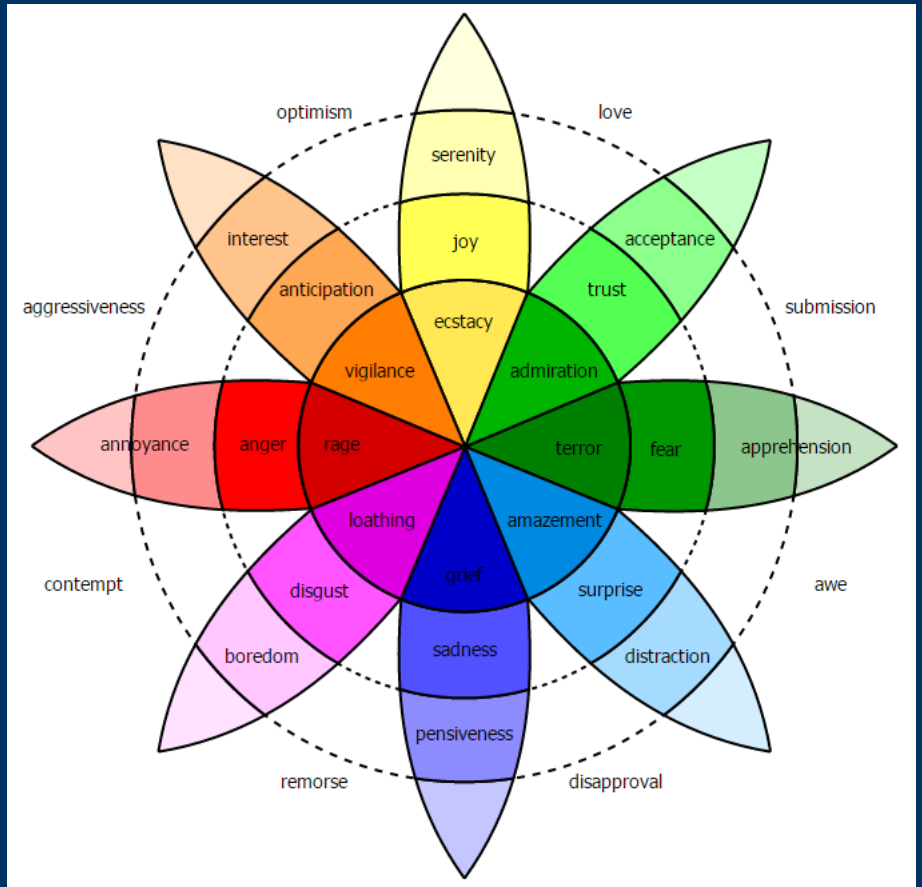
- ✓ conscious mental reaction (as anger or fear)
- ✓ subjectively experienced as strong feeling
- ✓ usually directed toward a specific object
- ✓ typically accompanied by physiological and behavioral changes in the body

Idea of emotions in ancient Greece

The conscious mind is the charioteer that tames the wild beasts that are emotions

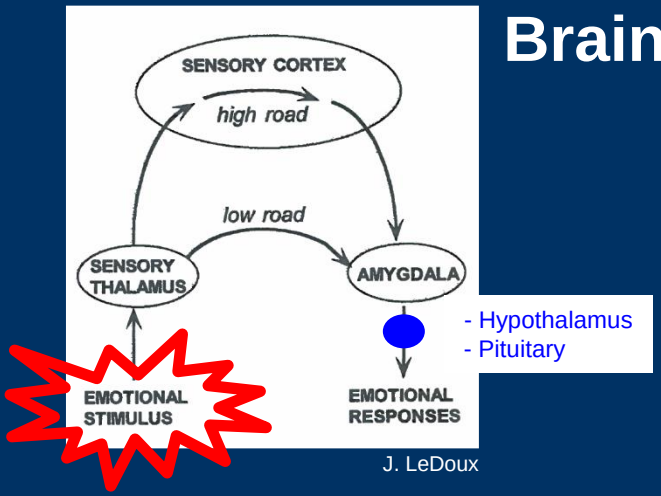
- Plato (428-348 BC)

Plutchik's wheel of emotions (1980)



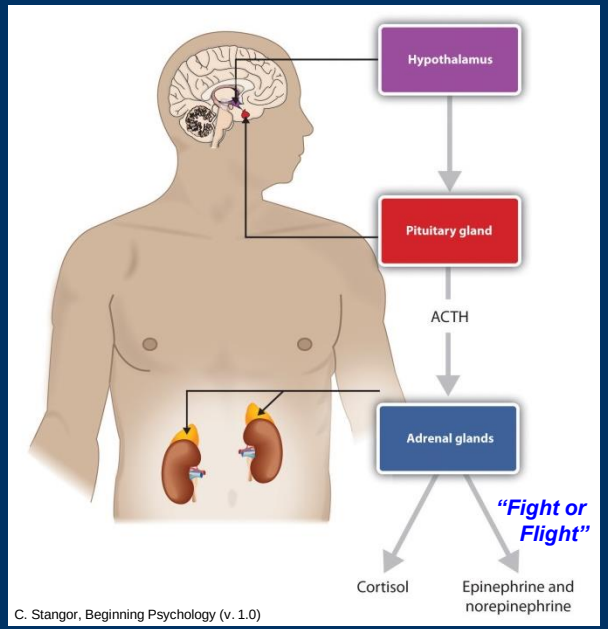
Anatomy of stress response

Brain



J. LeDoux

- Hypothalamus
- Pituitary



- Heart beats faster & blood pressure go up (more blood to the muscles, heart, etc.)
- Breathing rate increases & small airways in the lungs open wider for more oxygen
- Extra oxygen goes to the brain, increasing alertness
- Sight, hearing, and other senses become sharper
- Increase of blood sugar and fats for energy

Why study emotions like fear?

Anxiety (fear w/o external stimulus)

- **Panic**
- **Phobias**
- **Post-traumatic stress disorder**
- **Obsessive-compulsive disorder**
- **Generalized anxiety**

How are emotions like fear studied?

- Behavior
- Brain lesions (surgical in animals; accidental/stroke/cancer in humans)
- Brain stimulation
- Brain imaging
- Physiology & biochemistry

Little is known about genetics of emotions

- **Difficult due to complex genetics (humans are too diverse)**
- **Could reveal biochemical pathways for therapeutic targeting**

Why dogs?

High trait diversity

- morphological
- physiological
- disease
- *behavioral (incl fear & aggression)*

Low genetic diversity

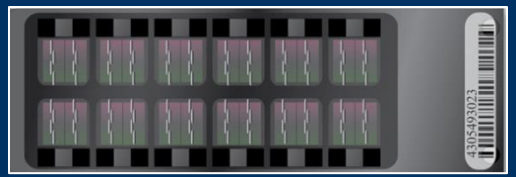
- 100-times simpler

Example: dog size varies 65X by breed; 250X by individuals

Ongoing studies to map dog behaviors

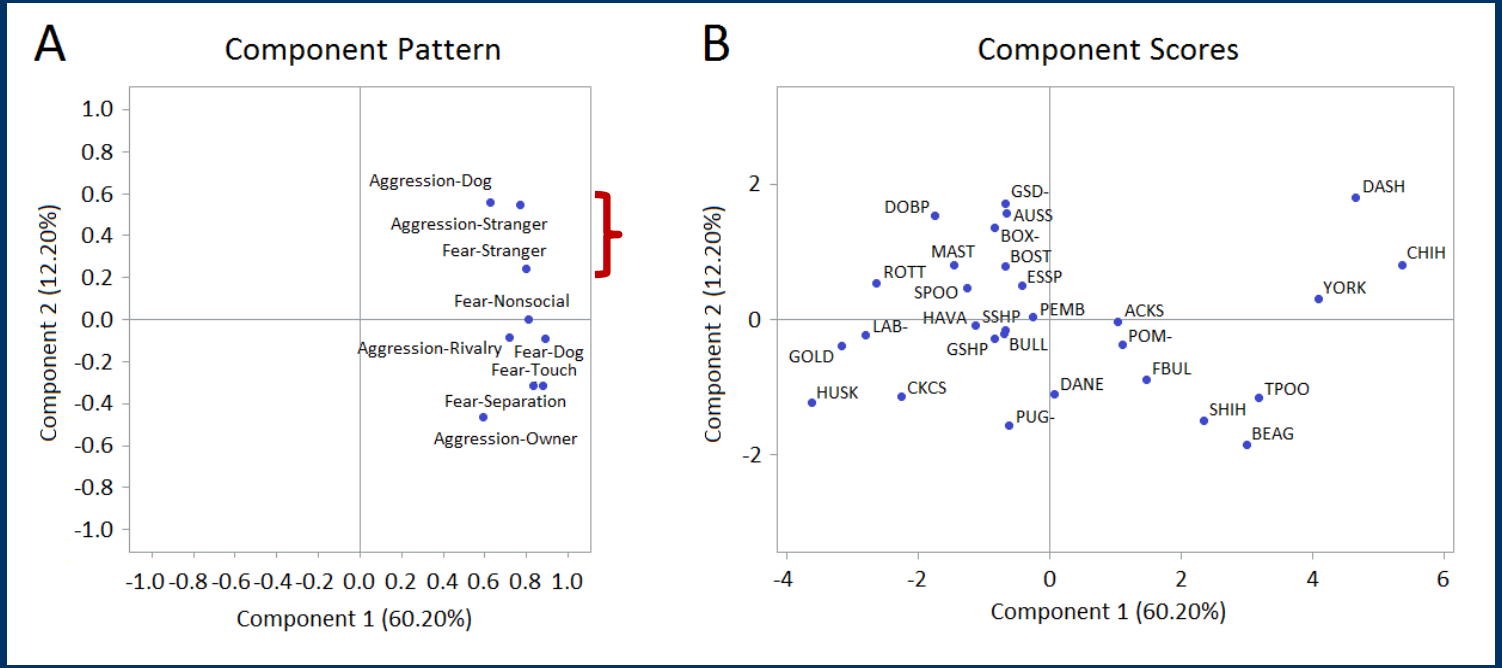
<u>Data 1:</u> dog owner survey of behavior across breeds (currently >40,000; http://vetapps.vet.upenn.edu/cbarq)

<u>Data 2:</u> genetic marker data for individual pure breed dogs (~1,200, >30 breeds)



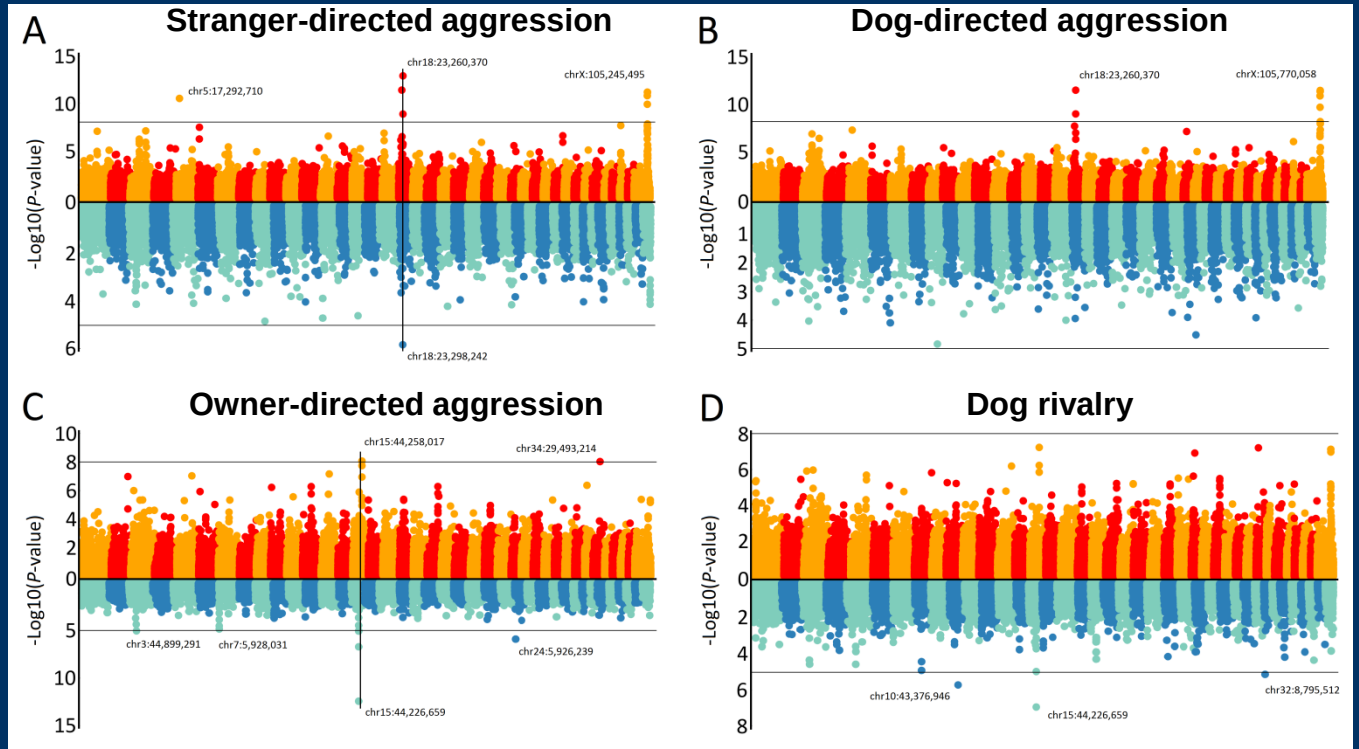
175,000 markers, 1/14,000 bp

Results 1: Principal components analysis



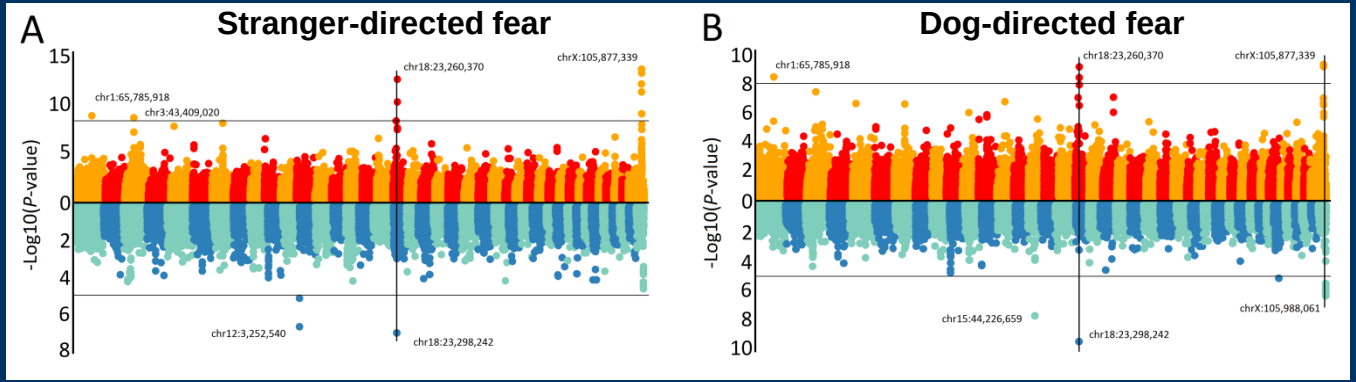
- Subset of fear & aggression cluster apart from other behaviors
- (A) Reanalysis of only fear & aggression traits suggests dog and stranger-directed aggression are related to fear of strangers … and these are distinct from owner-directed aggr.

Results 2: Genetic mapping of aggression



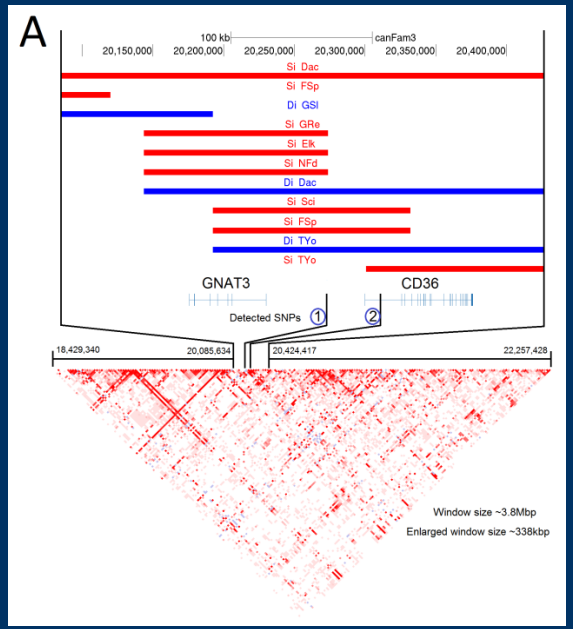
- Consistent with PCA, stranger and dog-directed aggression are genetically similar … and distinct from other aggression

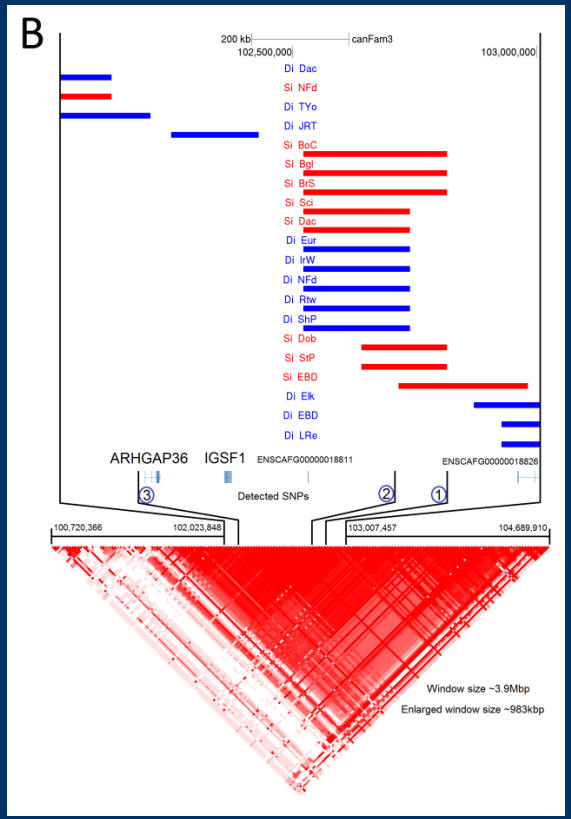
Results 2: Genetic mapping of fear



- Stranger & dog-directed fear are genetically similar to each other and to stranger & dog aggression
- Only those fear traits share both genome regions (vs. 3 others that have chr10 and either chr18 or chrX regions)

Results 3: precise mapping of genes

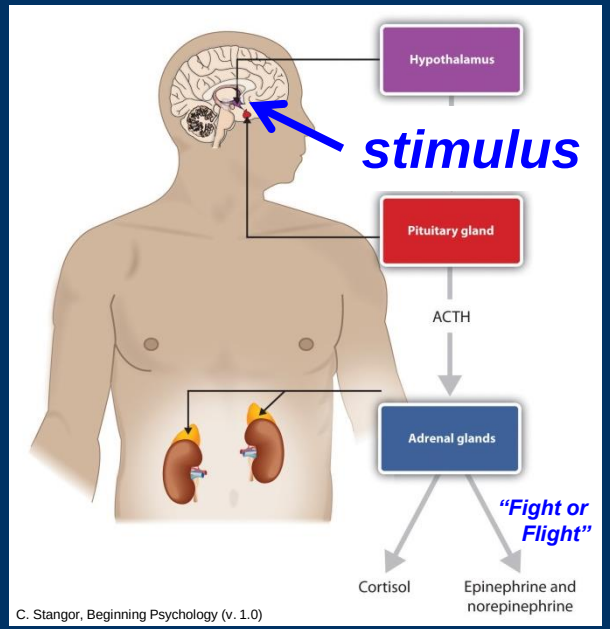




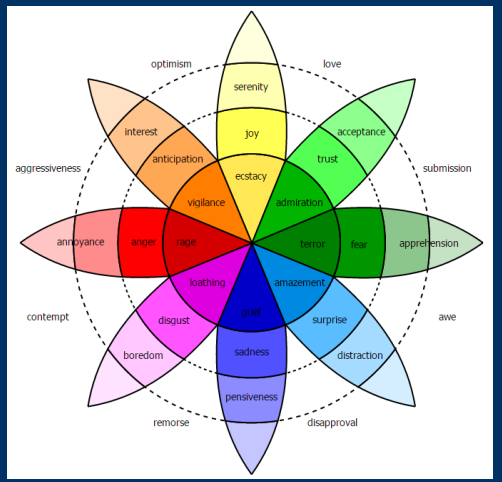
Fine-mapped by evolutionary selection

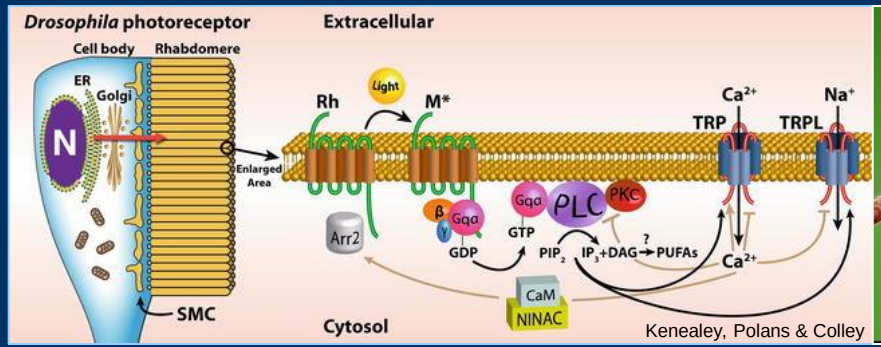
- chr18 gene as GNAT3
- chrX gene as IGSF1

GNAT3 is very biologically-relevant (1)

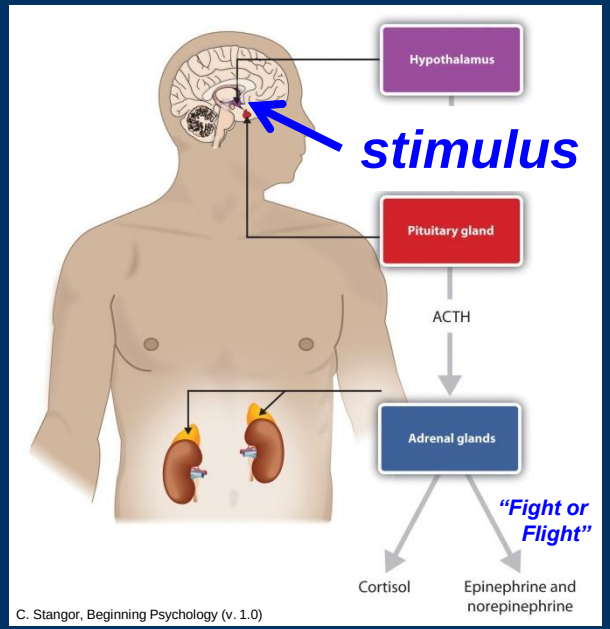


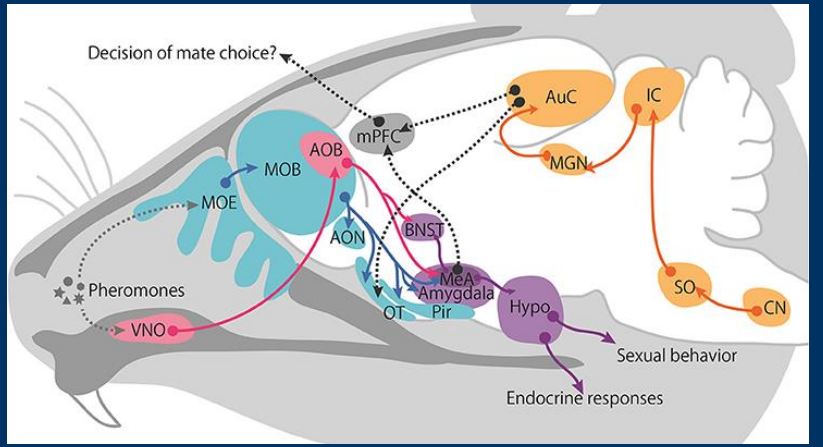
Chr18 *GNAT3* encodes "gustducin"
$G_q\alpha$ signal transducer for taste & pheromone receptors



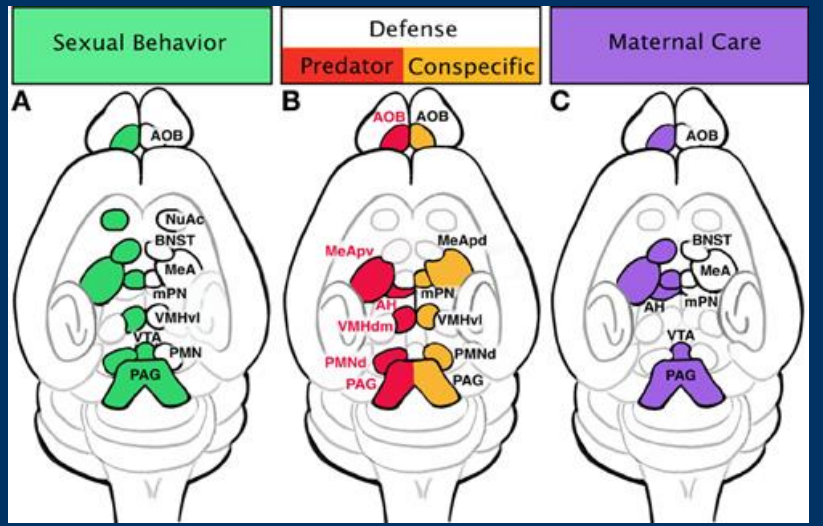


GNAT3 is very biologically-relevant (2)

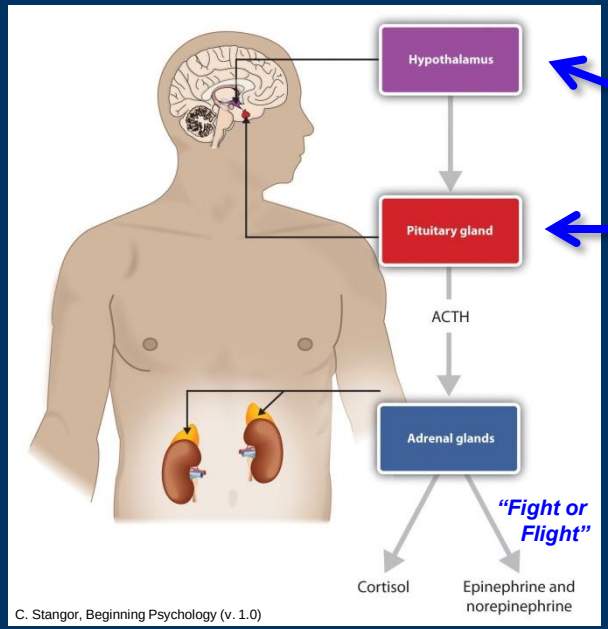




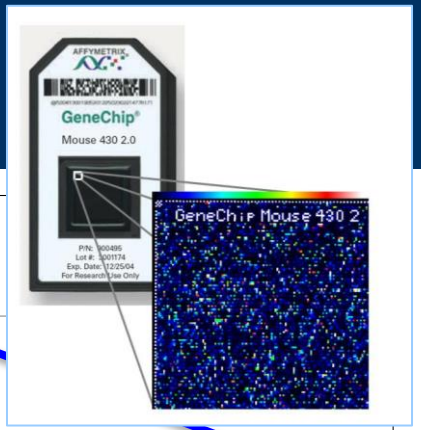
Chr18 GNAT3 encodes "gustducin" $G_q\alpha$ signal transducer for taste & pheromone receptors



IGSF1 is very biologically-relevant

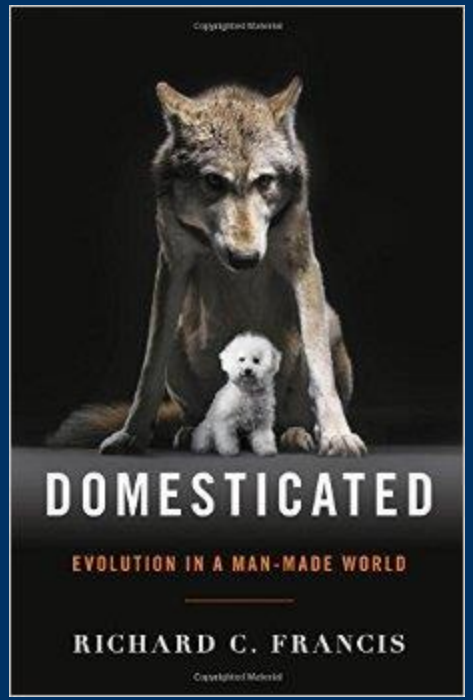


BioGPS



- ChrX *IGSF1* is expressed predominantly in pituitary & hypothalamus
- Human *IGSF1* mutation results in thyroid hormone deficiency

Frequency of variants & domestication



- Wolves seem to have high frequencies of fear-aggression variants
- Dogs have low frequencies

Conclusions

1. **The importance of emotions is underappreciated**
2. **Many aspects of life appeal to emotions: art, politics, *advertising*, etc.**
3. **We may often act irrationally due to emotions; *awareness* of this may mitigate damage**
4. **Fear & anxiety are major human problems**
5. **Dogs offer a genetic model of fear and aggression, which may sometimes be the same thing**

Appendix 12

Academic lecture/curriculum development:

OSU VME 6540 Structure & Function of Cells (Cell Biology), Lecture on canine genetics & genomics (Oct 5, 2015)

Genetic mapping & molecular dissection of dog traits

Carlos Alvarez, PhD

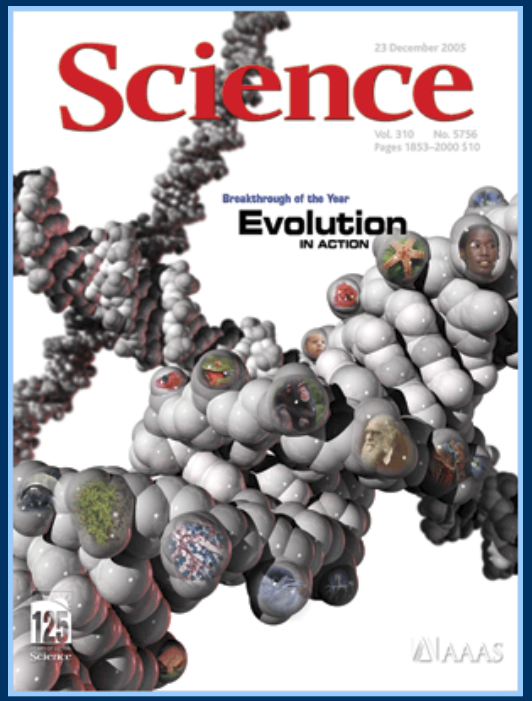
VME 6540
The Ohio State University College of Veterinary Medicine
October 5, 2015

Outline

1. Introduction
2. DNA copy number variation (CNV)
3. Genetic mapping of cancer risk

The Genome Era: historical context

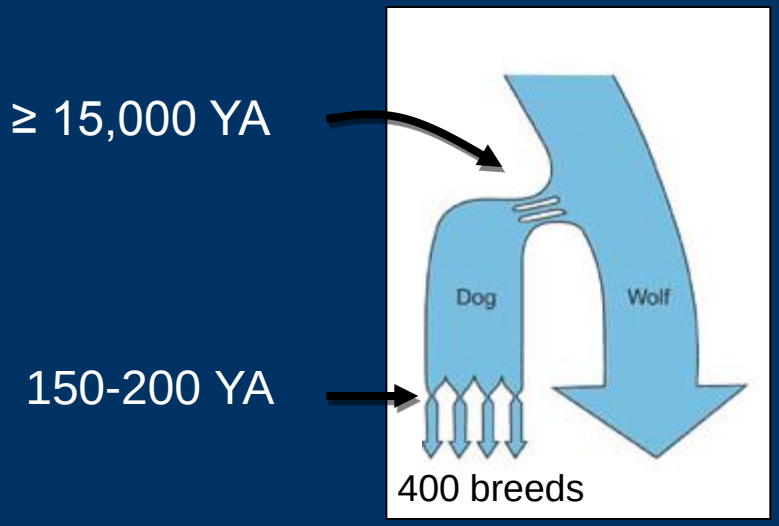
- 2000 Genome Sequencing
- 2005 Evolution in Action
- 2007 Human Genetic Variation



Genetics as a handle on biology & disease

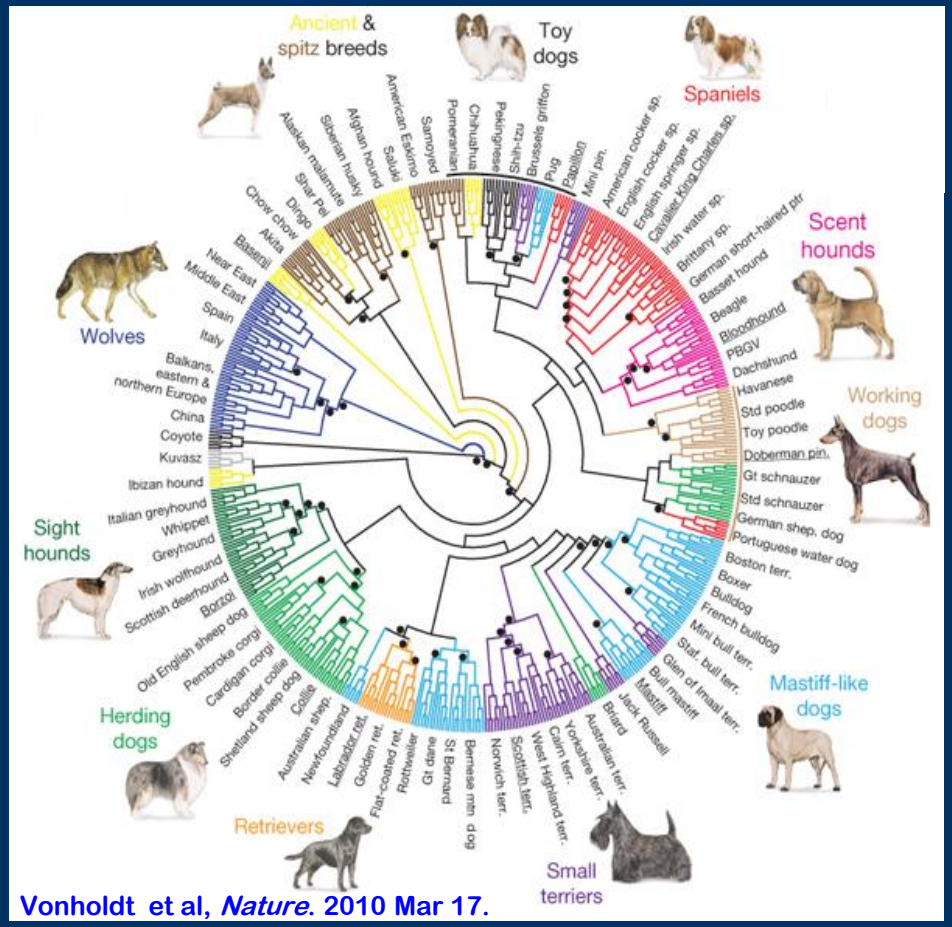
Why study CNV in dogs? (1)

Genetics



- 450 inherited diseases
- Dog genome similar to human
- 5-7X shorter lifespan v. humans
- 75M dogs in USA

- Intra-breed genetic variation low
- Inter-breed high (27 v 5-10% hum)

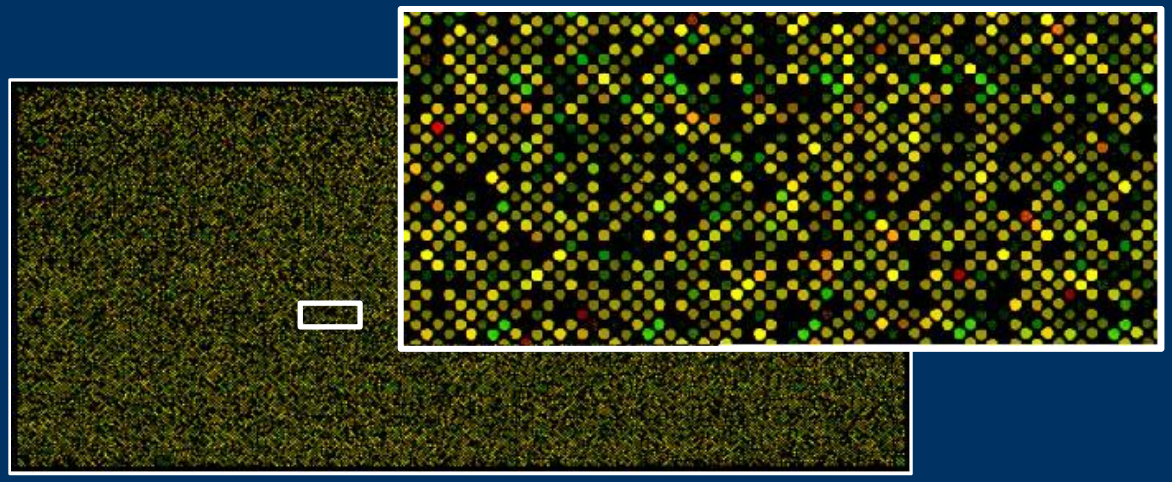


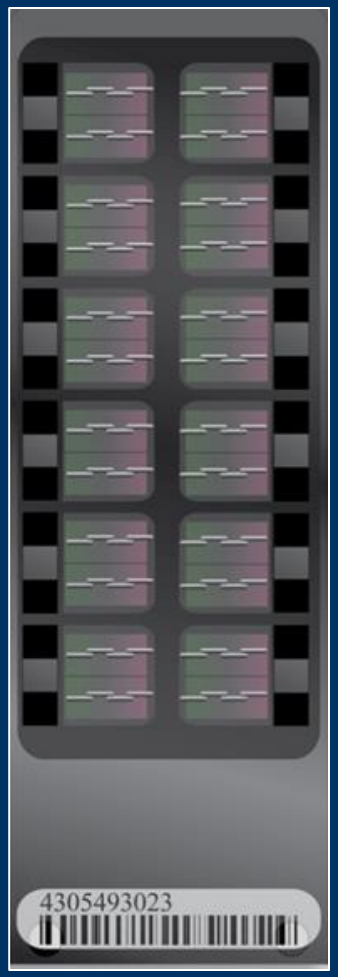
Why dogs? (2)

Variation
(morphological, physiological & behavioral)

- 65X difference sm/largest breeds (wt.)
- ~250X smallest/largest individuals

Dog genome sequence & tools



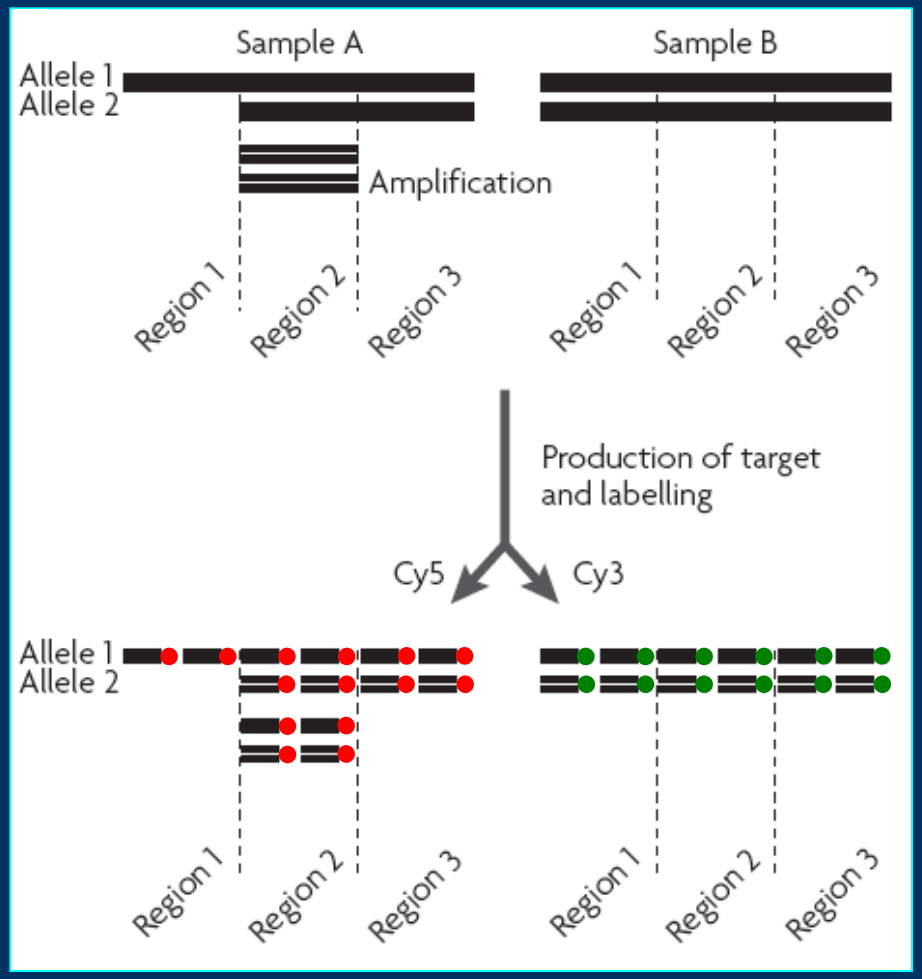


Functional variation: Custom 1,000,000 feature microarray (Agilent CGH)

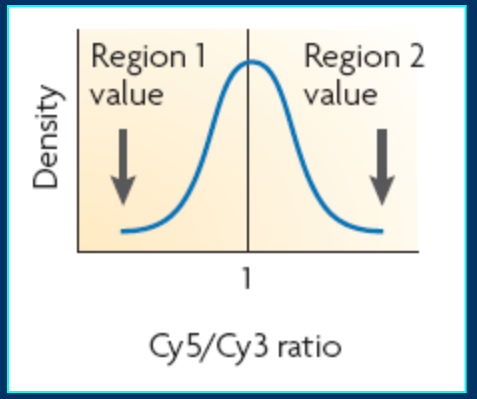
Genetic markers: 175,000 feature microarray (Illumina SNP)

… Emerging: High-throughput sequencing

Comparative Genomic Hybridization (CGH)



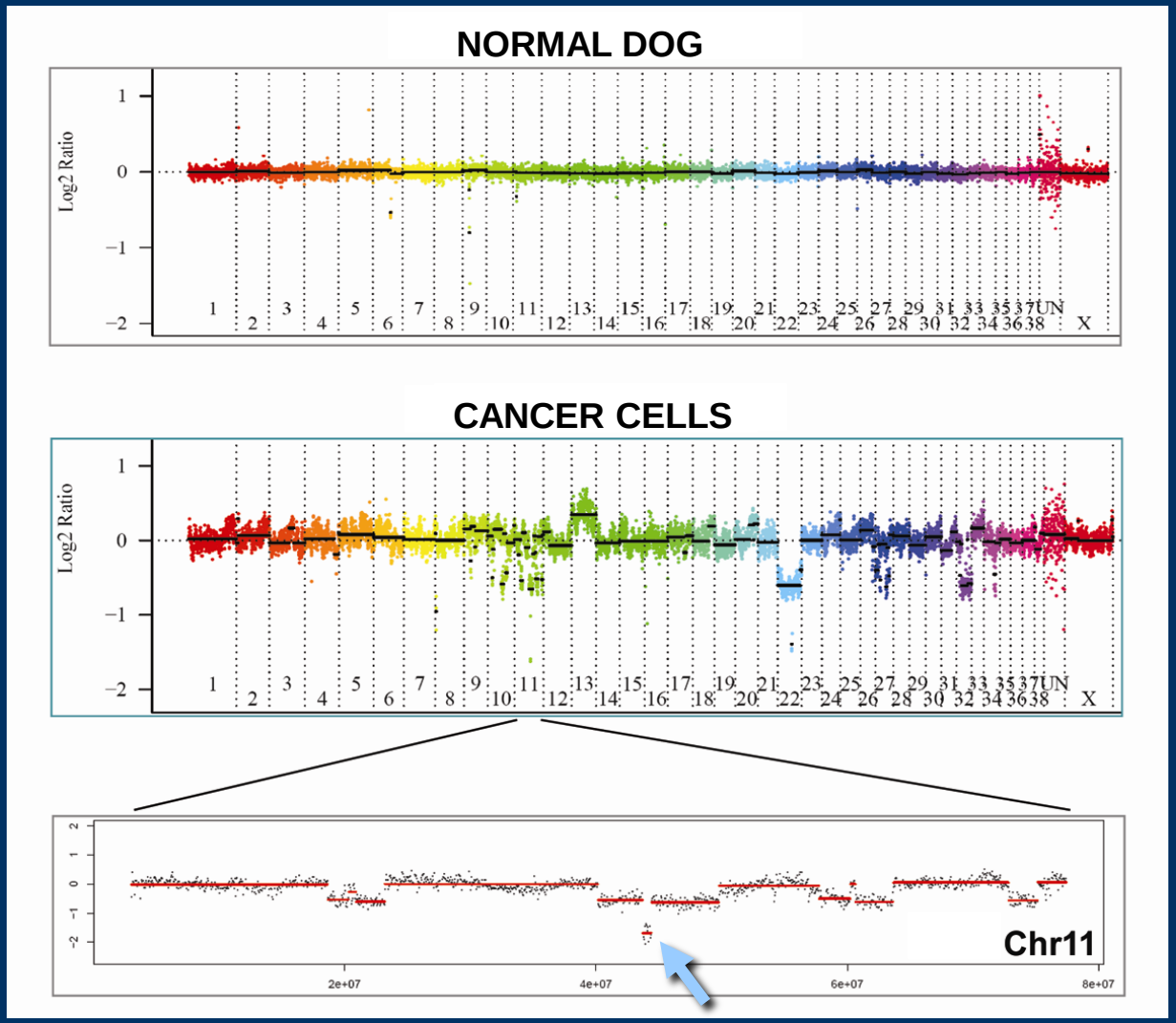
Hybridize

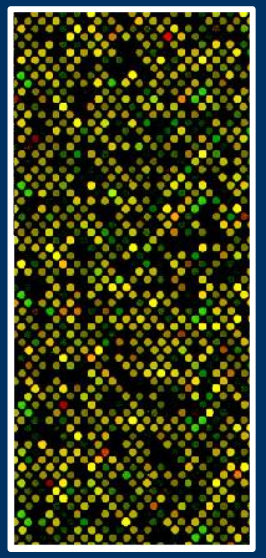


Greshan et al. Nature Reviews Genetics 9:291

2A. Identifying copy number *mutations* in cancer cells (*in somatic cell DNA*)

DNA copy number in cancer cells





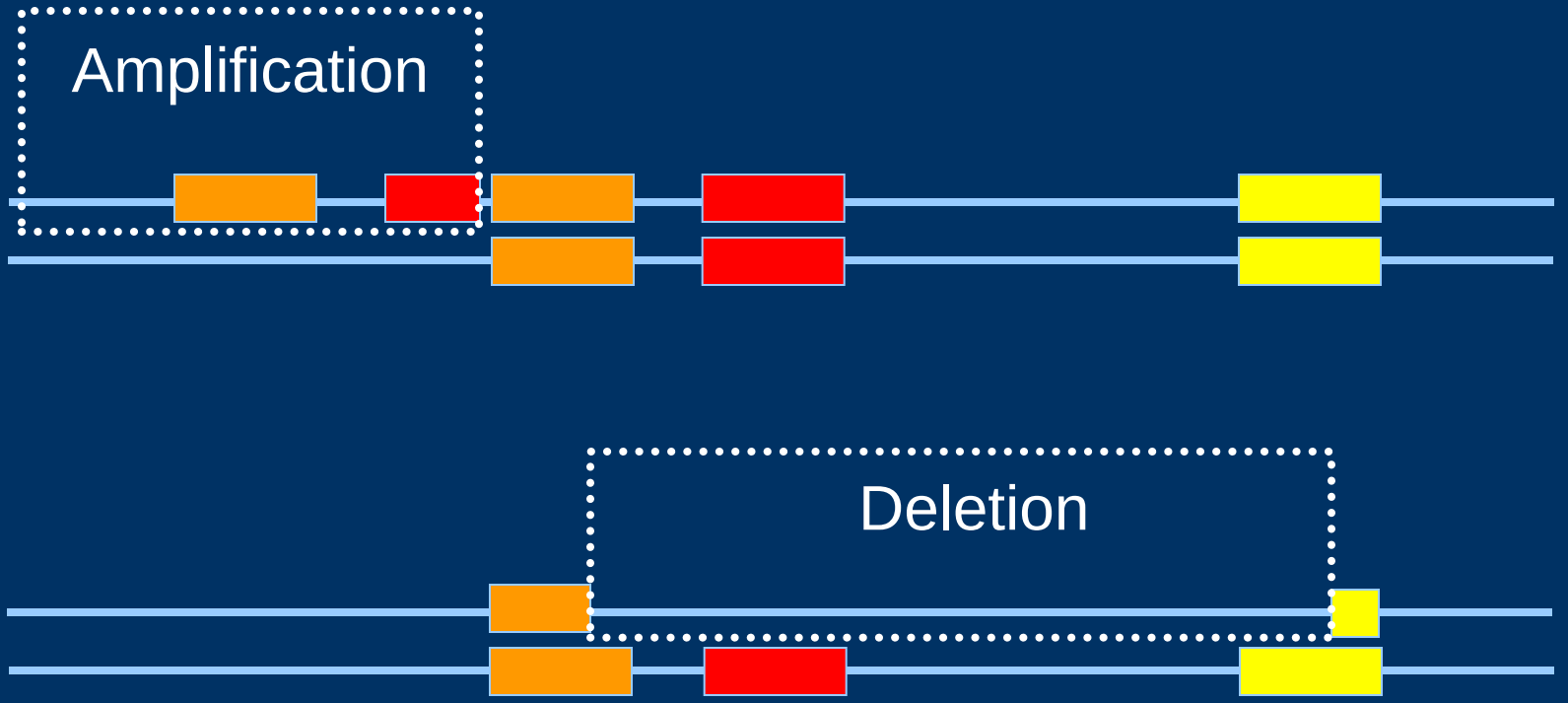
Comparative genomic hybridization

p16 tumor suppressor

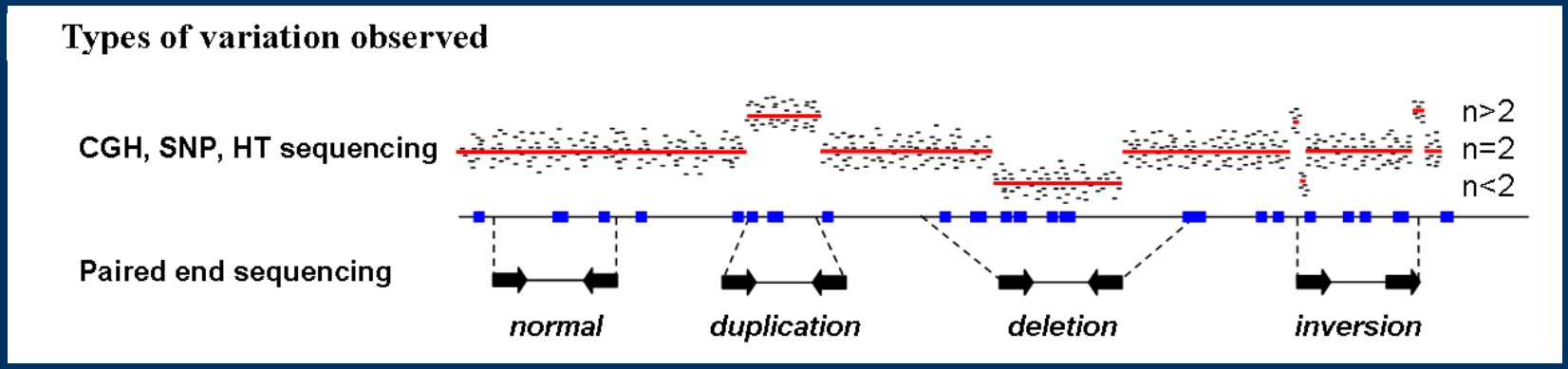
2B. Identifying inherited functional *variation* (in *inherited germ line DNA*)

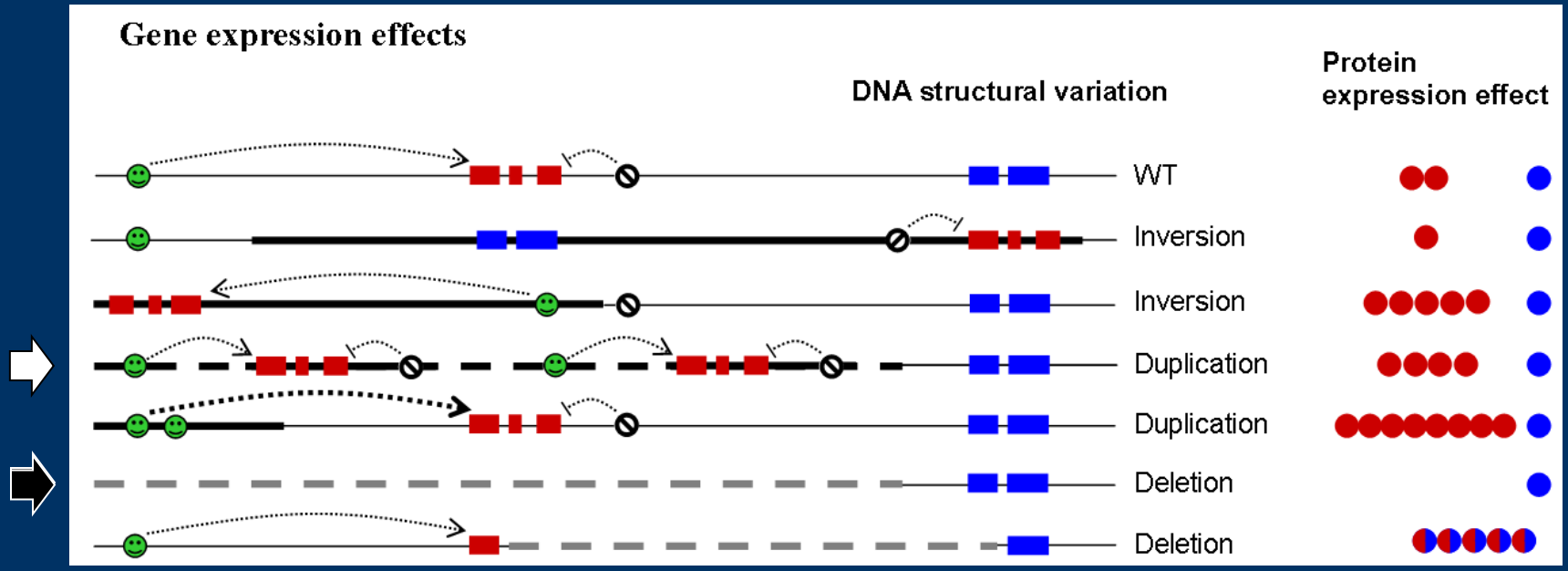
DNA copy no. variation (CNV)

Copy number alteration >50 bp
- *germ line* polymorphisms and mutations



CNV detection and effects





Alvarez & Akey, *Mamm Genome* 2012

Biological significance: rapid mutation mechanism in evolution

Liberles et al. Naturwissenschaften doi:10.1007/s00114-008-0446-0 (2008)

- Bats that feed on mammalian blood have 4 copies of a *Plasminogen Activator* gene

- Others, which feed on bird blood, have 1 copy

CNV in humans

- >>5000 CNVs described in humans
- ~800 CNVs/individual; ~150 affect genes
- Old, recurrent, and *de novo* events
- Overrepresented biology: senses, immunity, metabolism, microRNA, disease susceptibility
- Several disease-associations found

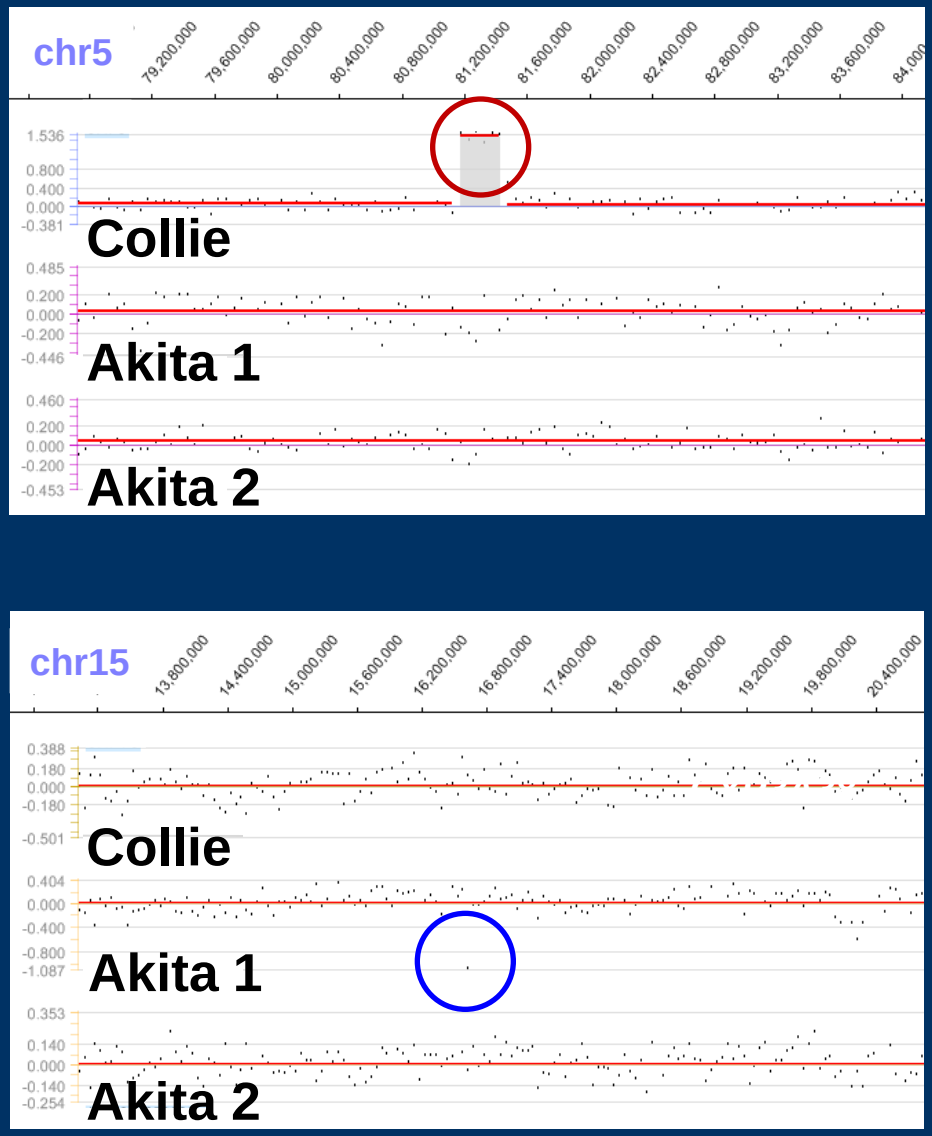
Inter-individual variation ~0.5% (by base pairs)
(vs. single nucleotide polymorphisms/SNPs, 0.1%)

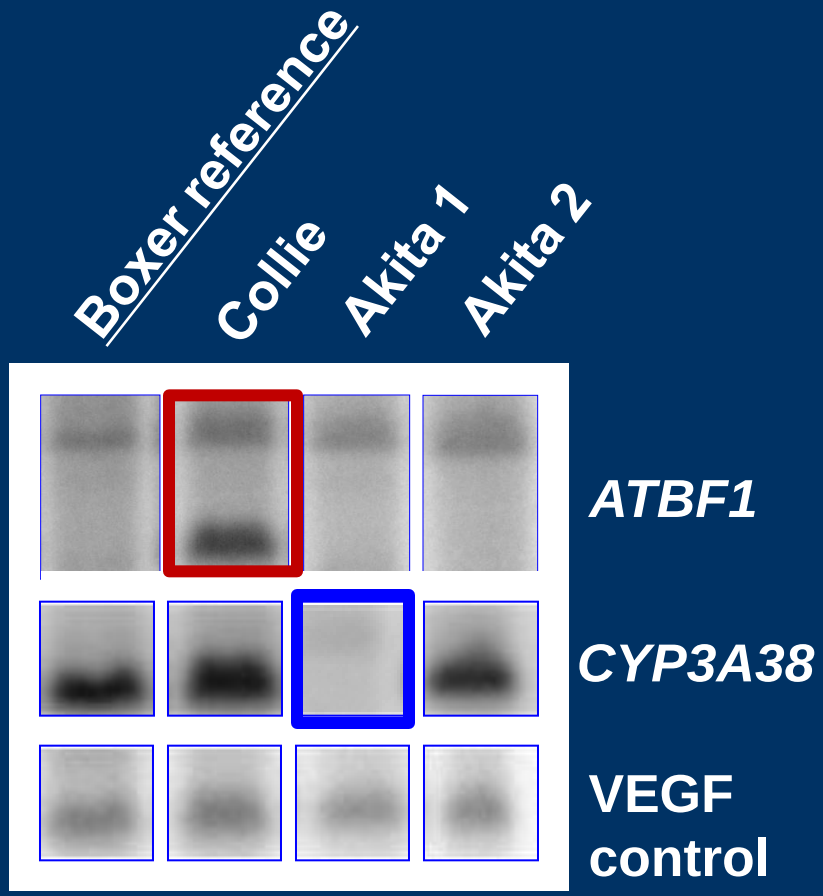
CNV mutations and dog traits, e.g.:

Trait (listed in order of variant size, SINE/LINE mechanism, CNVs with unknown effects)	D		me	Variant
Gonadal dysgenesis, X-chromosomal monosomy (presence of single sex chromosome, X)	D			Loss of sex chromosome (normal complement of autosomes and a single sex chromosome, X, or 77,X0)
* Hemophilia A	Ir			Large inversion
* Skeletal myopathy and dilated cardiomyopathy (Duchenne muscular dystrophy)	G			> 2.4 Mb deletion
Hemophilia, "hemophilia B"	L			Deletion
Cone degeneration	A			>140 kb deletion
Copper toxicosis	B			39.7 kb deletion
Hemophilia, "hemophilia B", factor IX deficiency	P			>30 kb deletion
Cone-rod dystrophy, "locus 3"	G			>20 kb deletion
Paroxysmal hypertonicity, "episodic falling syndrome (EFS)"	C			15.7kb deletion
Collie eye anomaly	R S L			7.8 kb deletion
Oculoskeletal dysplasia	S			1,267 bp deletion
* Cone-rod dystrophy	S			180 bp deletion
* Hair ridge ("Ridgeback") and predisposition to dermoid sinus	R			133 kb duplication
* Chondrodysplasia	Basset Hound, Cairn Terrier, Cardigan Welsh Corgi, Chihuahua, Dachshund, Dandie Dinmont Terrier, Glen of Imaal Terrier, Grand Basset Griffon Vendeen, Havanese, Japanese Chin, Lancashire Heeler, Norwich Terrier, Pekingese, Pembroke Welsh Corgi, Petit Basset Griffon Vendeen, Scottish Terrier, Shih Tzu, Skye Terrier, Swedish Valhund, Tibetan Spaniel, West Highland White Terrier, Yorkshire Terrier	18	5 kb insertion; retrotransposition of FGF4 retrogene
* Furnishings (mustache and eyebrows)	Australian Terrier, Briard, Brussels Griffon, Cairn Terrier, Giant Schnauzer, Glen of Imaal Terrier, Havanese, Irish Wolfhound, Norwich Terrier, Old English Sheepdog, Portuguese Water Dog, Scottish Deerhound, Scottish Terrier, Shih Tzu, Standard Poodle, Standard Schnauzer, Toy Poodle, West Highland White Terrier, Yorkshire Terrier	13	167 bp insertion
Globoid cell leukodystrophy	Irish setter	8	78 bp insertion
Cone-rod dystrophy, "locus 1"	Miniature Longhaired Dachshund	15	44 bp insertion

Alvarez & Akey, *Mamm Genome* 2012

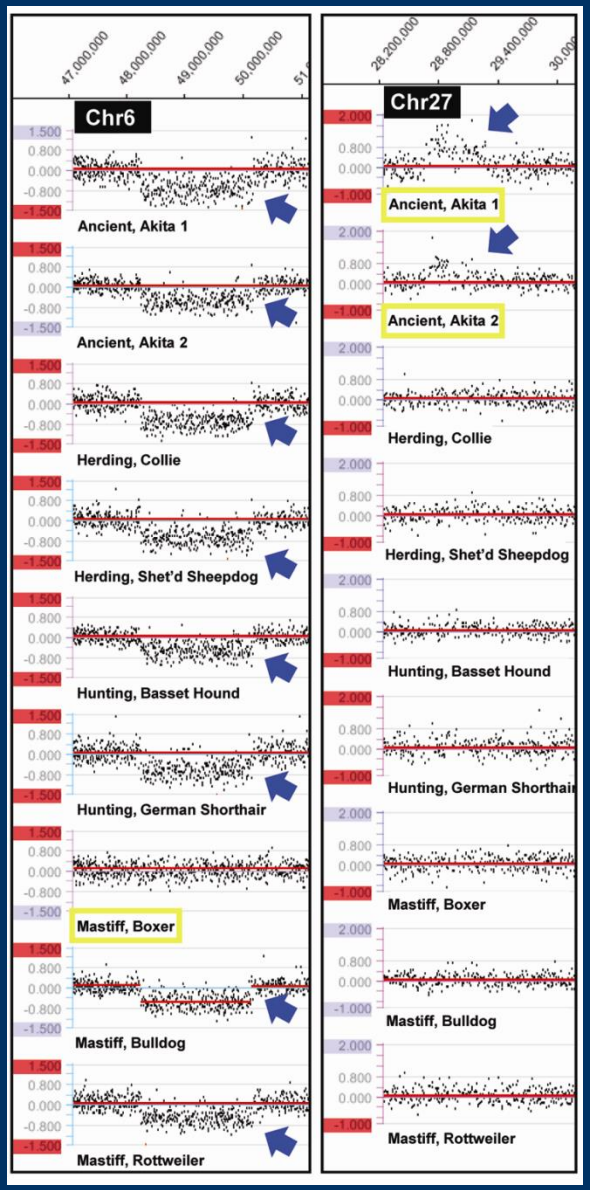
DNA structural / copy no. variation





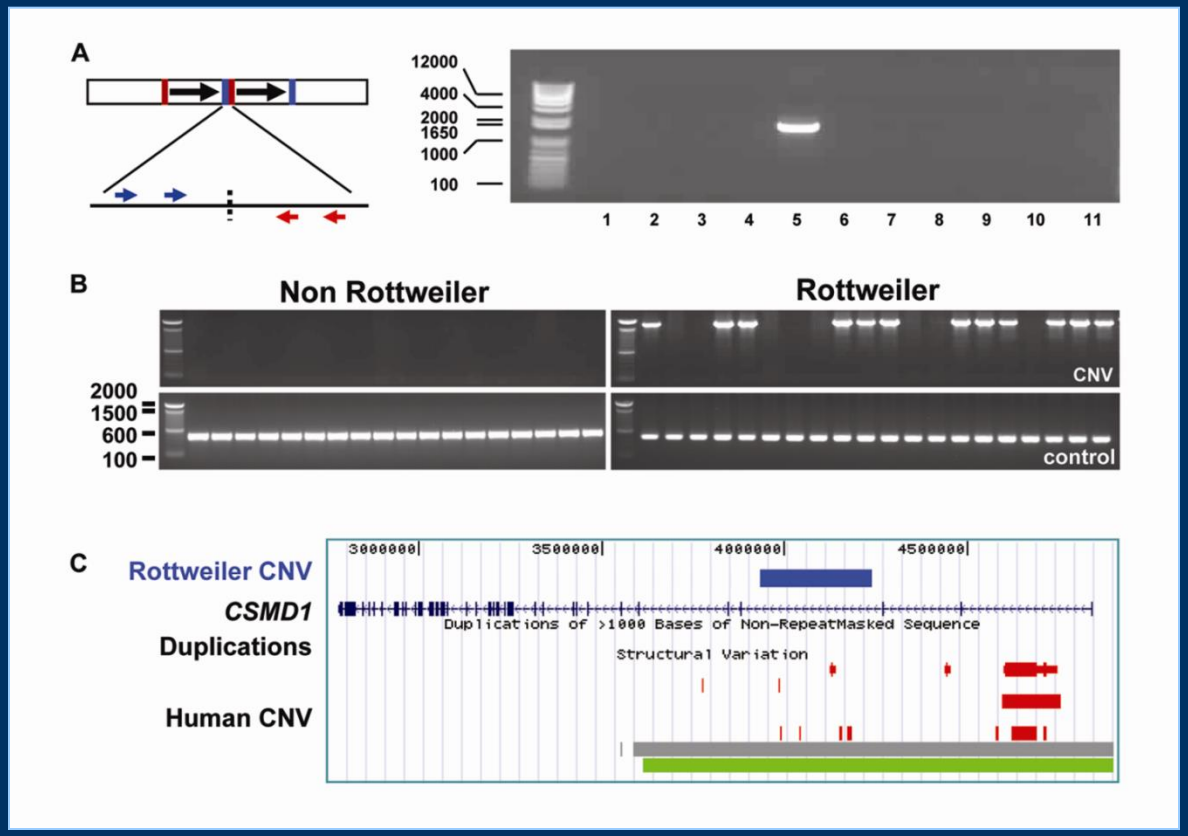
aCGH of normal dogs

Candidate breed-specific CNVs



CNV in *CSMD1* of Rottweilers

- *CSMD1* frequently deleted in epithelial tumors
- Somatic point mut. & allele loss associated with cancer
- Deletion and reduced expression assoc. w/ poor outcome



Dog CNV: numbers & affected genes

Genome Research 2009

Resource

Mapping DNA structural variation in dogs

Wei-Kang Chen,[1,4] Joshua D. Swartz,[1,4,5] Laura J. Rush,[2] and Carlos E. Alvarez[1,3,6]

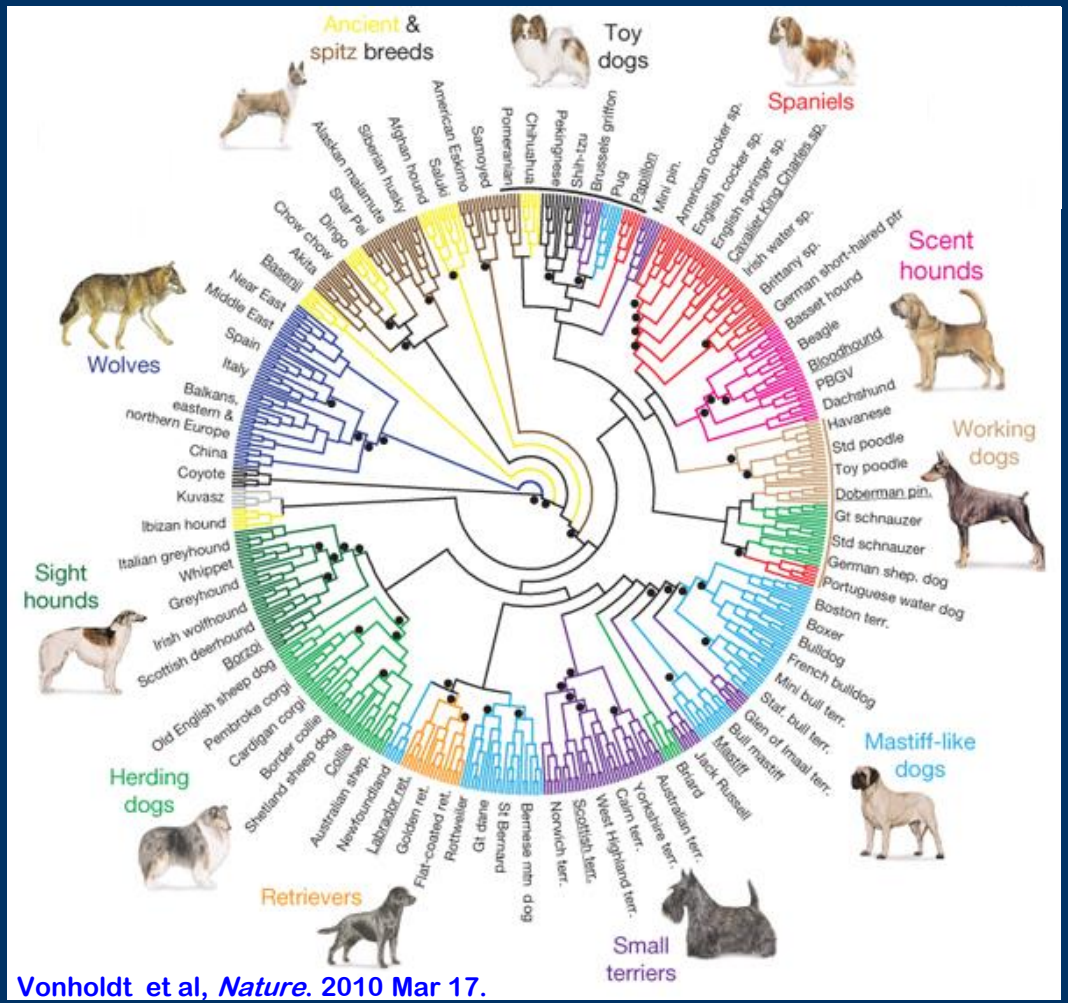
[1]*Center for Molecular and Human Genetics, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio 43205, USA;* [2]*Department of Veterinary Biosciences, The Ohio State University, Columbus, Ohio 43210, USA;* [3]*Department of Pediatrics, The Ohio State University College of Medicine, Columbus, Ohio 43210, USA*

- Identified ~200 CNVs (80 CNVrs); 75 genes
- Dogs/rodents/humans have similar numbers
- Several human disease genes affected
- Similar gene functions are over-represented (e.g., immunity and olfaction)

Breeds of interest

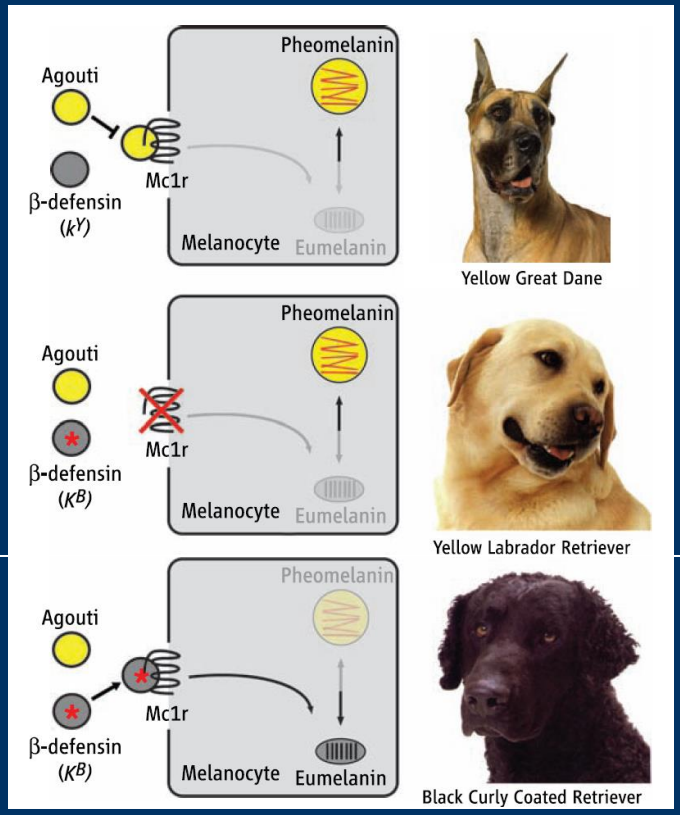
Driven by veterinary investigators

- Cavalier King Charles Spaniel
- German Shorthaired Pointer
- Labrador Retriever
- Doberman Pinscher
- Great Dane
- Shar-Pei
- Shih-Tzu
- Cocker Spaniel
- German Shepherd
- Beagle
- Border Collie
- Yorkshire Terrier
- Boxer
- Bulldog
- Greyhound



Vonholdt et al, *Nature*. 2010 Mar 17.

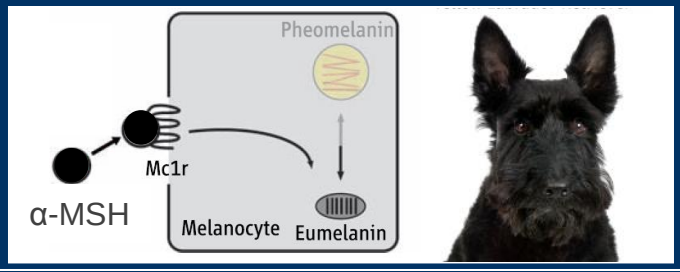
Functional variation: coat color



Candille et al. (2008) *Science* 318

β-defensin 3 = *Dog* Dominant Black (K^B)

3 *Mammalian* coat color genes: Agouti, MC1R, Tyrosinase



BASIC COLOURS	
A (agouti) = *agouti signalling protein* (***ASIP***) CFA24	amongst reddish hairs in some breeds)
a^y	Fawn/sable (cream to yellow to red with darker tips) (some solid black hairs intermingled
a^w	Wolf sable – wild type colour (many banded hairs – black-reddish-black)
a^t	Black-and-tan or brown-and-tan
a	Recessive black
B (brown) = *tyrosinase related protein 1* (***TYRP1***) CFA11	
B	Black eumelanin
b (***b*** ^{s}, ***b*** ^{d}, ***b*** ^{c})	Brown eumelanin
E (extension) = *melanocortin receptor 1* (***MC1R***) CFA5	
E ^{M}	Melanistic mask
E	Eumelanin (black, brown, blue) can be produced
e	Only phaeomelanin (red, yellow, cream) produced
K (from 'dominant blacK') = (***CBD103***) CFA16	
K^B	Black, brown or blue (eumelanin pigmentation only)
k^{br}	Brindle (on body region that would be phaeomelanin pigmented otherwise)
k^y	Expression of agouti alleles that express phaeomelanin possible

Schmutz & Berryere (2007) *Animal Genetics* 38:539-549.

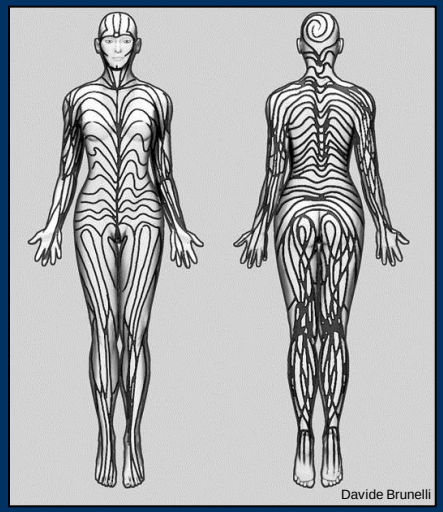
Dominant Black dominance order

K^{Black} > $k^{brindle}$ > k^{yellow}

Kerns et al. (2007) Genetics 176:1679.

Brindle is similar to Blaschko's lines

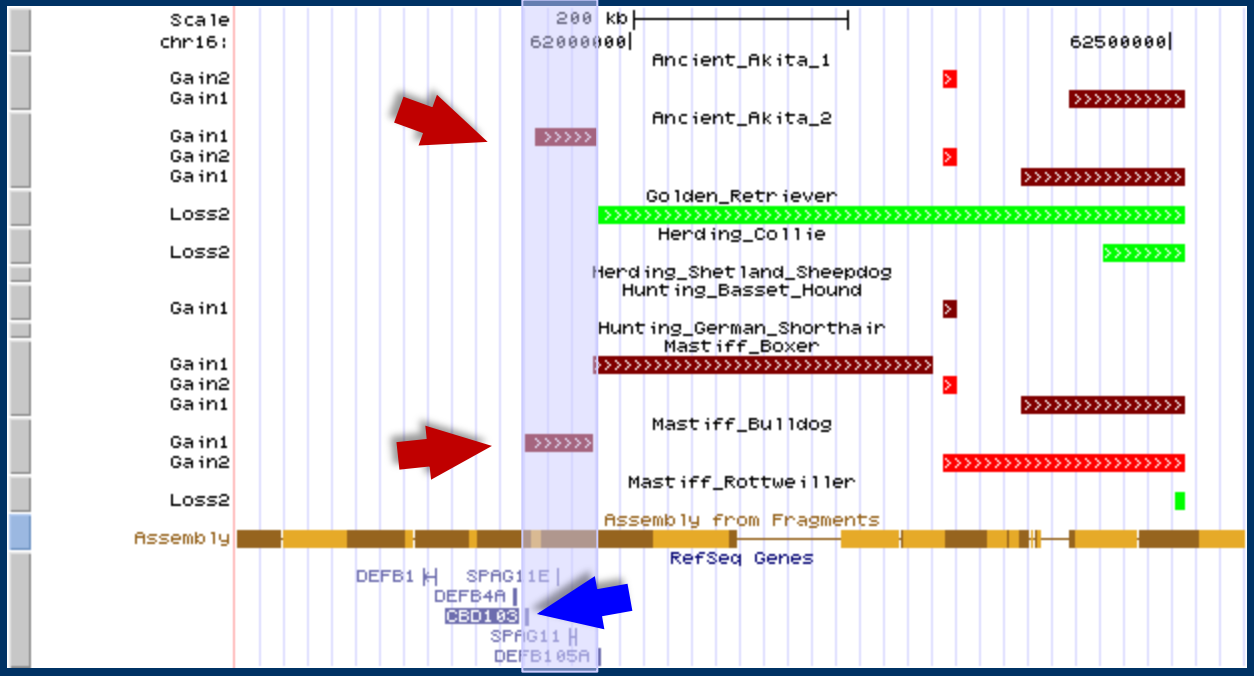
Pigmentary mosaicism (not mutation, but effect of two genetic cell types)



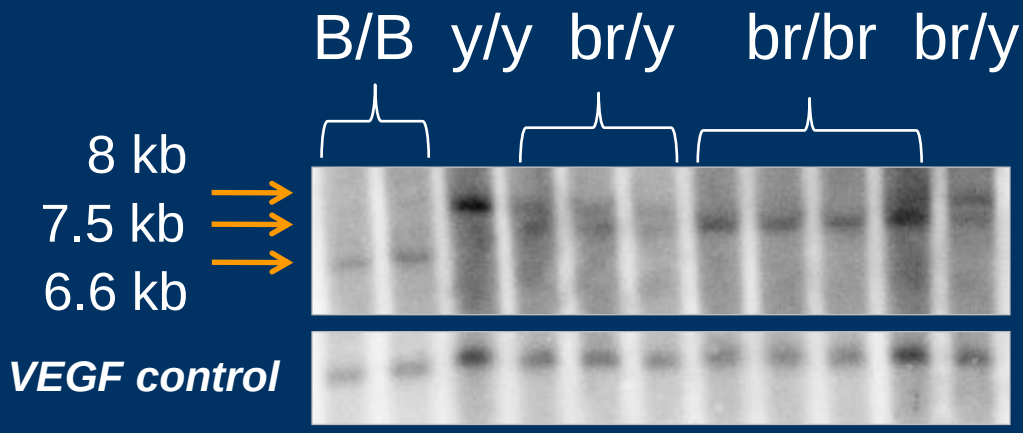
Brindle:
Suggestive of a DNA mutation with an *epigenetic* effect

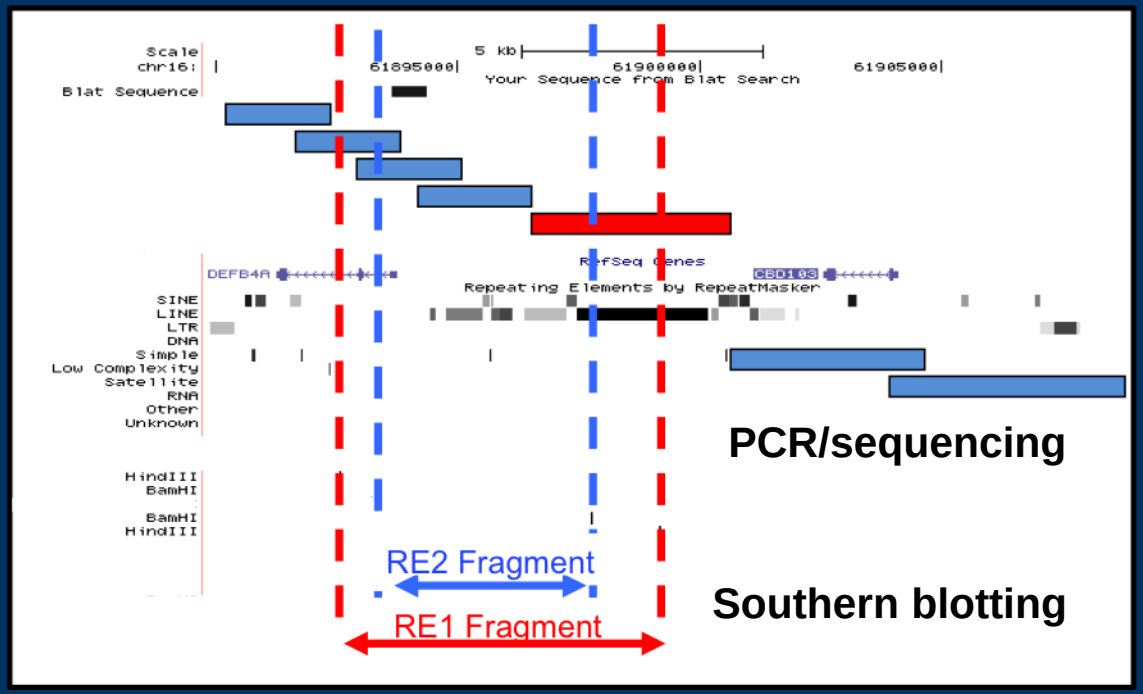
Kerns et al. (2007) *Genetics* 176:1679.

Brindle mutation of *Dom. Black*

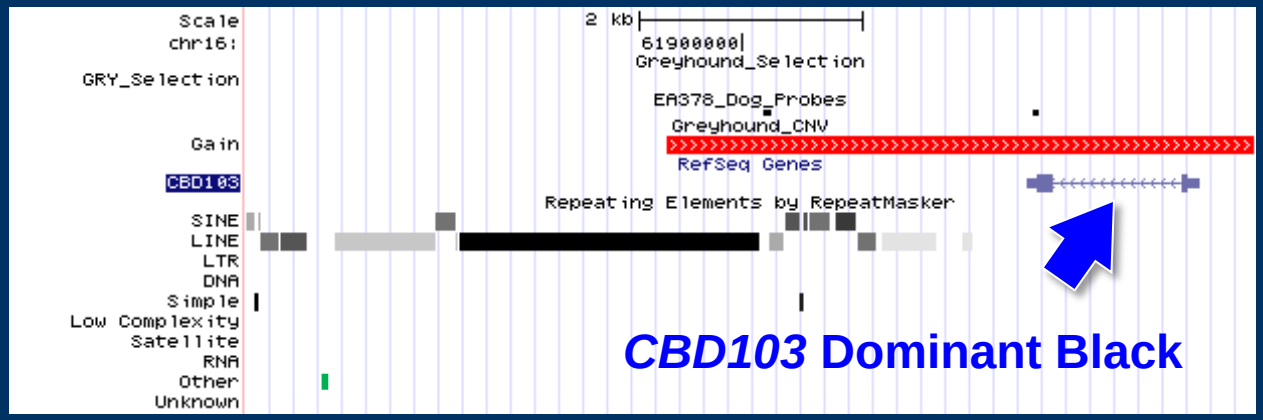


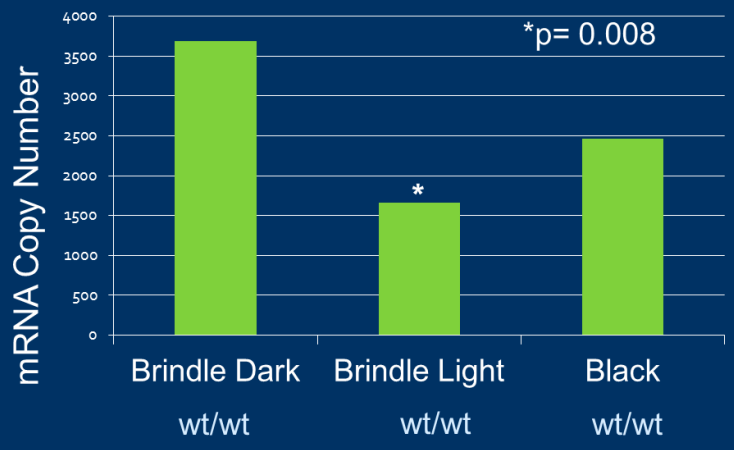
Characterization of K^{Black} alleles



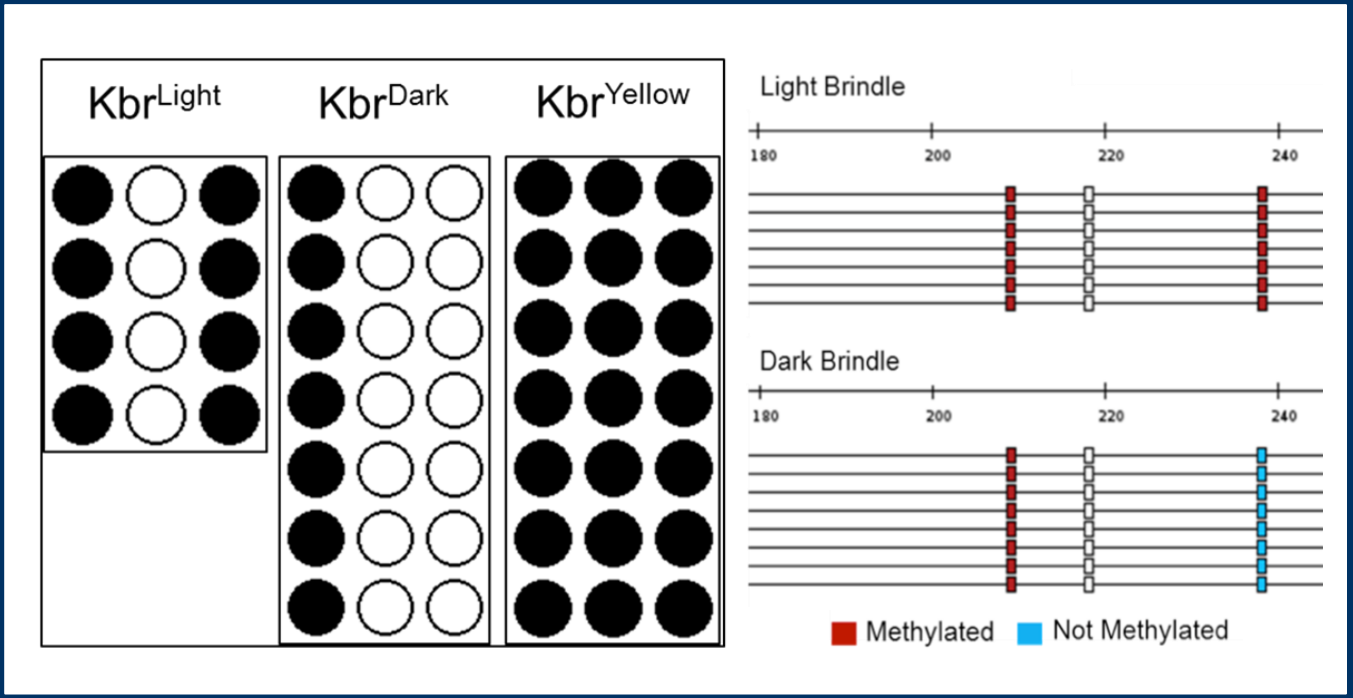


Brindle due to *CBD103* expression

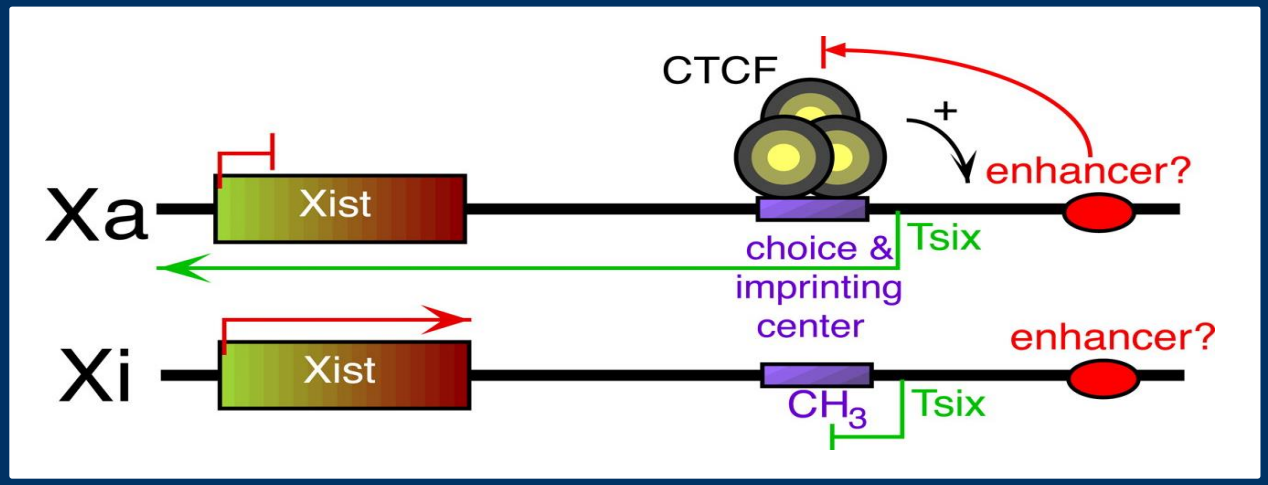




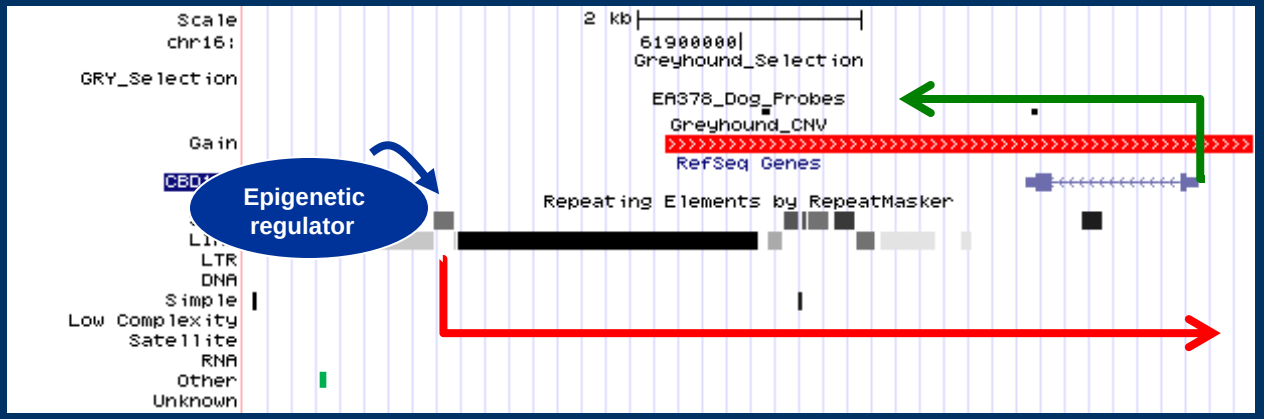
DNA methylation at CTCF locus



Brindle may be like X-inactivation



Chao et al. (2002) *Science* 295: 345-347.



3. Mapping complex disease traits: osteosarcoma risk

Greyhounds

- European nobility used for coursing
- AKC show, since 1885
- Racer registration, since 1906
- Increased risk for osteosarcoma *in racer* (17X RR)

Show

Racer

Disease	Description
Polyneuropathy	Juvenile dogs show exercise intolerance and walking difficulties. In the later stages, there is severe muscle atrophy and ataxia
Pancreatic acinar atrophy	Diabetes accompanied by diarrhea
Thyroid function	Basal serum thyoxine (T4) and free thyroxine (fT4) concentrations were significantly lower in Greyhounds than in non-Greyhounds
Vitreal degeneration	Degeneration of the vitreous, often associated with cataract
Sodium thiopentane sensitivity	Sensitivity to sodium thiopentane and other barbiturate anaesthetics
Neoplasia - Osteosarcoma	Excess of osteosarcoma
Persistent right aortic arch	Aortic arch abnormally producing an encircling ring around trachea and oesophagus.clinical sigs of regurgitation, aspiration pneumonia and failure to thrive due to oesophageal compression
Asthma	Asthma, broncho-constriction - non specific bronchial hyperresponsiveness
Meningoencephalitis	Head tilting, ataxia, recumbency, circling, and blindness. Severe inflammatory changes in caudate nucleus, cerebellum and brain stem
Cutaneous vasculopathy	Skin and less often kidney lesions: hemorrhages, fibrinoid arteritis, thrombosis with deep, slowly healing ulcers of skin. Renal glomerular necrosis
Epilepsy	Fitting and seizures
Malignant hyperthermia	Hyperthermia and rigid paralysis after exercise, stress or exposure to halothane, caffeine .
Sensorineural deafness	Deafness starting at roughly 4 weeks
Type 1 von Willebrand disease	Blood clotting disorder
Progressive Retinal Atrophy	Loss of night vision progressing to total blindness
Hindlimb lameness	Lameness caused by avulsion of the tibial tuberosity
Chronic superficial keratitis of the cornea	Bilateral disease of the cornea -vascularised pigmented subepithelial growth

Other Greyhound traits

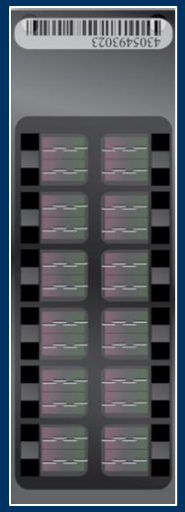
Running performance

Sleep

Single nucleotide polymorphisms (SNPs)

```
          *
...ACCTGCTATTTCGATT...
   |||||||| ||||||||
...ACCTGCTAGTTCGATT...
          *
```

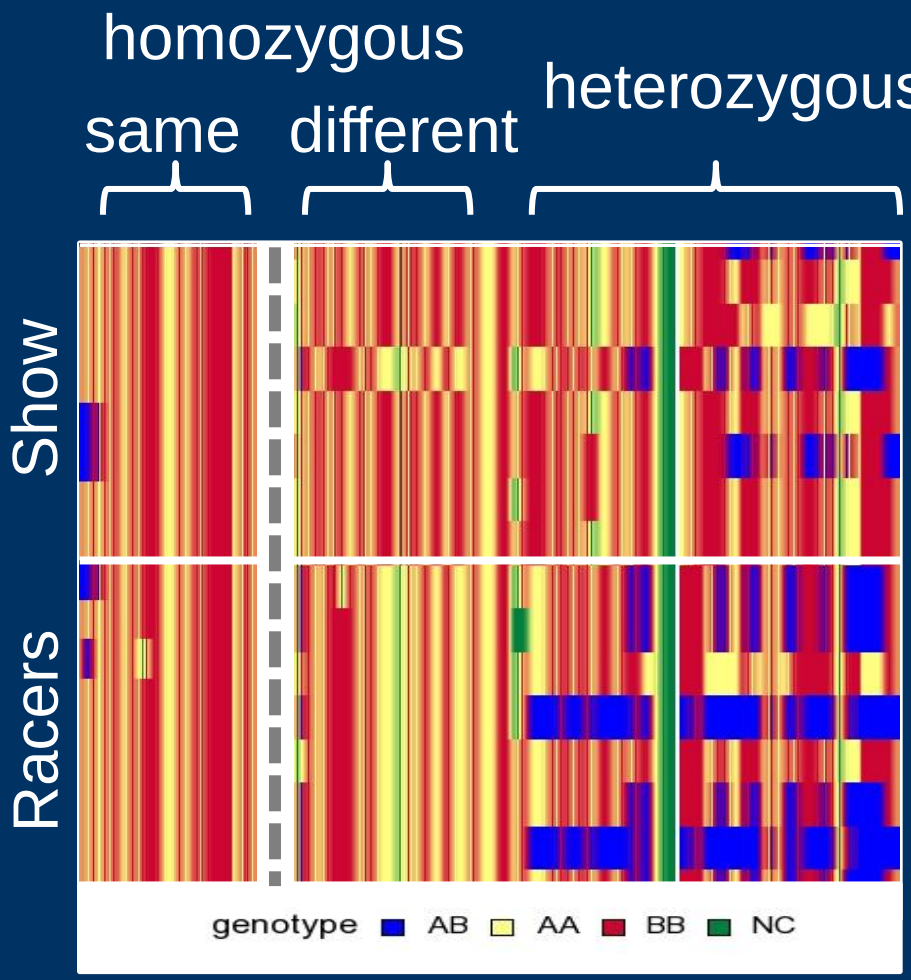
Genetic markers for *mapping* (usu. not causative mutations)



Illumina SNP array:

- ~175,000 SNPs (=1 SNPs/14,000 bp)
- Avg call rate: 99.8%; Reproducibil: >99.9%

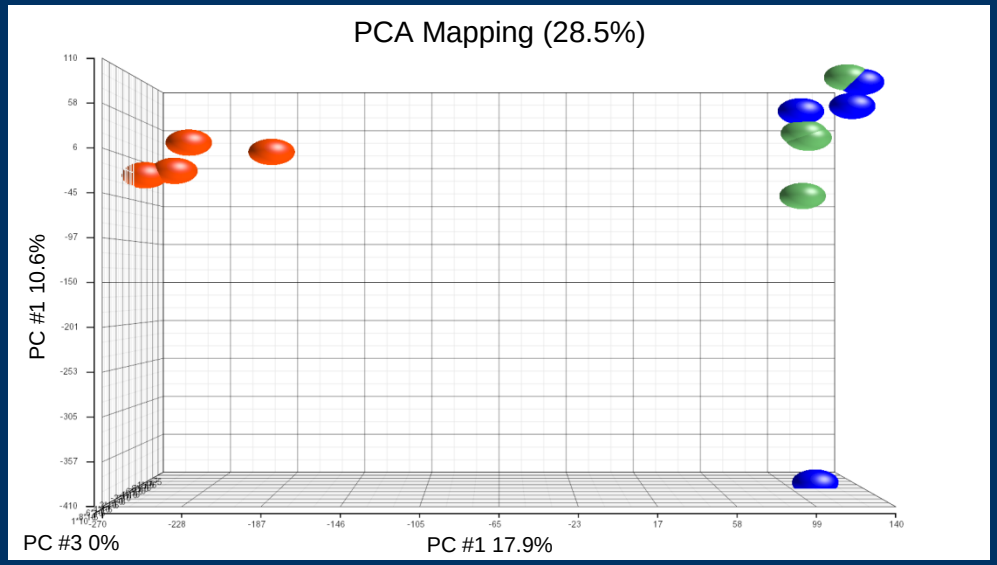
Comparing racers and show



SNP genotypes

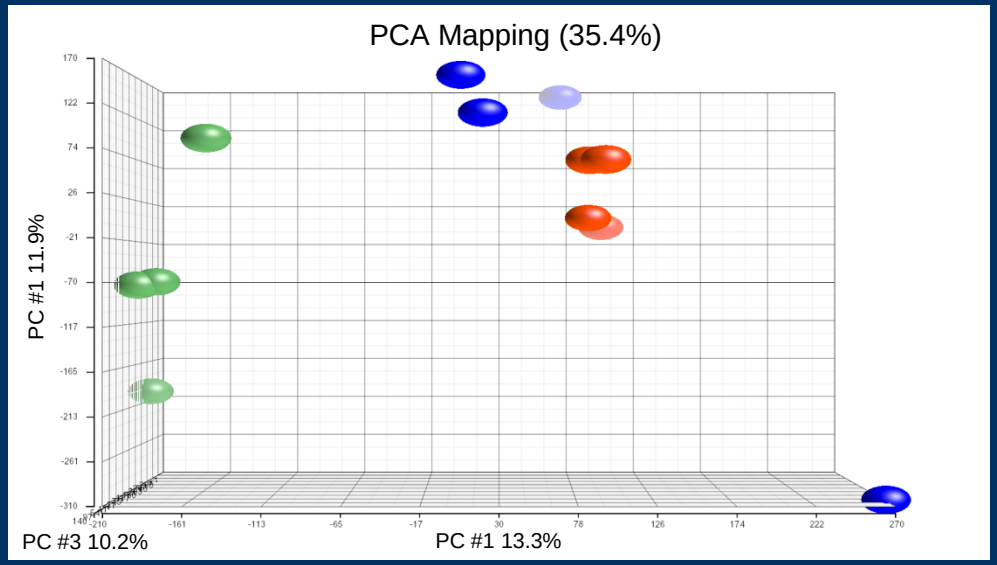
Grouping dogs by OSA status

- OSA+/- racers group apart from shows



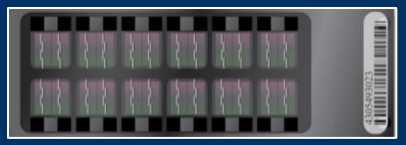
OSA+ Racer
OSA - Racer
Show

- Removing “racer” effect, OSA- racer and shows group together



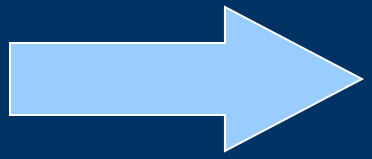
Genome wide association (GWA)

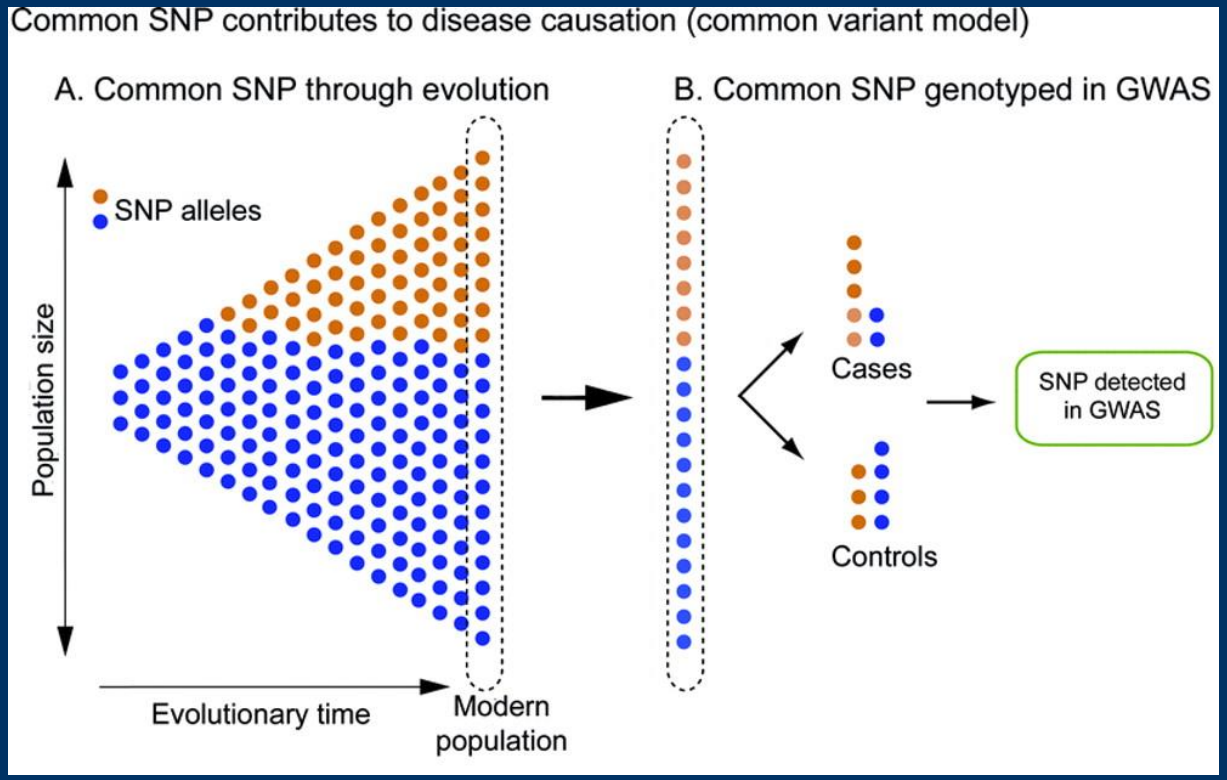
SNP microarray:
175,000 SNPs



Frequency of ea SNP allele (A or B) is measured in cases vs. controls

Example of *single* SNP





Which SNP alleles over whole genome have different frequency in cases vs. controls?

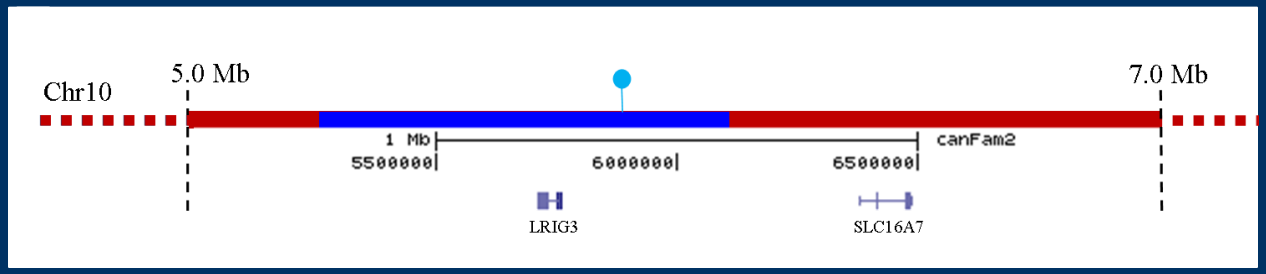
Osteosarcoma genome wide association (GWA) & Results

1. **Case control study:** **12 OSA+ racers**
   12 OSA - racers
   12 OSA - show

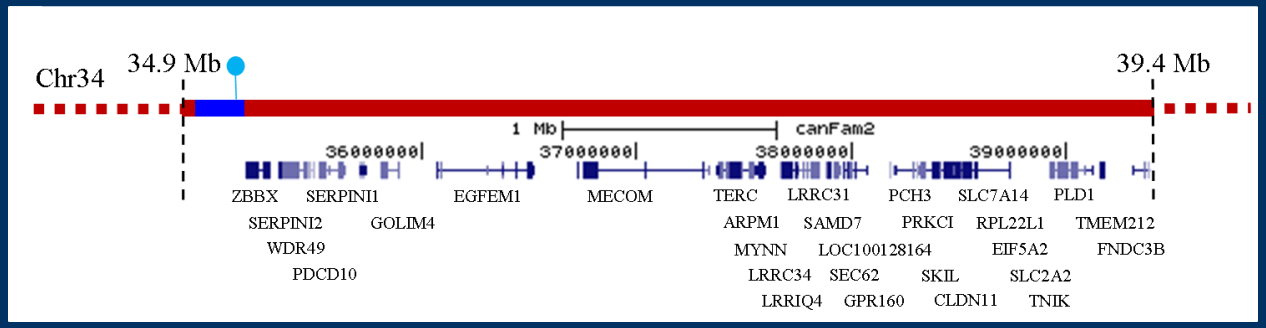
2. **Second group of case-control racers for validation**

- **Mapped complex genetic OSA risk loci: *15 hits***
- ***Validated 3*** **OSA loci in a second group of dogs**

GWA-mapped loci



OSA1



OSA2

Chr16 **Rex1/Zfp42** (undifferentiated ES cell marker)

OSA3

LRIG3 locus under selection

Tracking footprints of artificial selection in the dog genome

Joshua M. Akey[a,1], Alison L. Ruhe[b], Dayna T. Akey[a], Aaron K. Wong[b], Caitlin F. Connelly[a], Jennifer Madeoy[a], Thomas J. Nicholas[a], and Mark W. Neff[b,c,d,1]

Table S3. Candidate selection regions with single-gene resolution

Chr	Coordinates, Mb	Breed	Gene	Putative phenotype
4	82.12–83.12	Greyhound	*CDH9*	Behavioral
5	80.92–81.92	Greyhound	*ZFHX3*	Cranial morphology
7	51.14–52.14	Brittany	*PIK3C3*	Behavioral
9	11.16–12.06	Greyhound, Brittany	*SOX9*	Skeletal morphology
10	5.19–6.19	Greyhound	*LRIG3*	Elongated body axis
18	26.65–27.65	Dachshund	*SEMA3D*	Morphology
21	17.71–18.71	Poodle	*DLG2*	Behavioral
23	27.43–28.43	Border Collie	*SATB1*	Athletic performance

Conclusions

- Validated approach of cataloguing CNV in dogs
- Identified structural variation affecting 100s of genes (e.g., Dominant Black in brindle)
- Mapped and validated osteosarcoma risk loci
- Racing selection may increase OSA risk

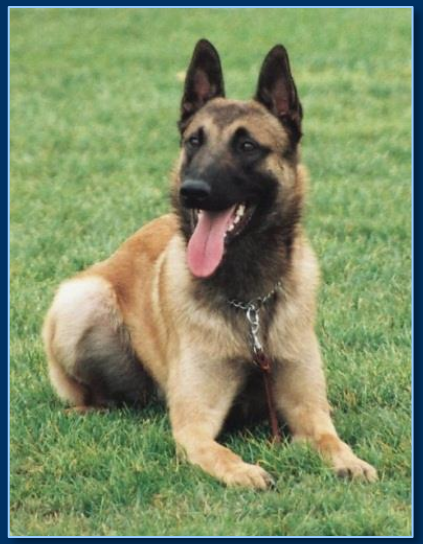
Next steps & ongoing studies

Next steps: identifying osteosarcoma-risk mutations, human relevance, translation

---

Ongoing:

- Canine lung and soft tissue sarcoma
- Longitudinal cohorts: cancer risk in DoD military dogs
- Determining canine Hb genetics
- Mapping behavioral traits



Thanks to the contributors

Nationwide Children's

Carlos Alvarez* Lab
> Jennie Rowell**
> Isain Zapata
> Leszek Rybaczyk
> Joshua Swartz
> Elise Fiala
> Michelle Perez
> Wei-Kang Chen
> Fu-Chuan Hsieh
> Isaac Wolfe

Grant funding
*NHGRI, DoD, AKC, MAF
**NINR

Participants & donors

OSU College of Veterinary Med.

Guillermo Couto
> Sara Zaldivar-Lopez
> Liliana Marin
> Cristina Iazbik

Lynette Cole
Wendy Lorch Lab

Cheryl London Lab
> Nicole Stingle, Clin Trials Office
> Tamra Mathie, Clin Trials Office

Laura Rush
Bill Kisseberth Lab
Ronaldo da Costa

Kun Huang, OSU BMI
> Terry Camerlengo
> Jacob Aaronson

Pramod & Bogdan Pathak, MSU