Measurement of segregating behaviors in experimental silver fox pedigrees

AV Kukekova[1,*], **LN Trut**[2], **K Chase**[3], **DV Shepeleva**[2], **AV Vladimirova**[2], **AV Kharlamova**[2], **IN Oskina**[2], **A Stepika**[2], **S Klebanov**[4], **HN Erb**[5], and **GM Acland**[1]

1*James A. Baker Institute for Animal Health, Cornell University, Ithaca, NY 14853, USA*

2*Institute of Cytology and Genetics of the Russian Academy of Sciences, Novosibirsk, 630090, Russia*

3*University of Utah, Department of Biology, Salt Lake City, Utah 84112, USA*

4 *New York Obesity Research Center, St. Luke's-Roosevelt Hospital Center, New York, NY 10025, USA*

5*Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, NY 14853, USA*

Abstract

Strains of silver foxes, selectively bred at the Institute of Cytology and Genetics of the Russian Academy of Sciences, are a well established, novel model for studying the genetic basis of behavior, and the processes involved in canine domestication. Here we describe a method to measure fox behavior as quantitative phenotypes which distinguish populations and resegregate in experimental pedigrees. We defined 50 binary observations that nonredundantly and accurately distinguished behaviors in reference populations and cross-bred pedigrees. Principal-component analysis dissected out the independent elements underlying these behaviors. PC1 accounted for >44% of the total variance in measured traits. This system clearly distinguished tame foxes from aggressive and wildtype foxes. F1 foxes yield intermediate values that extend into the ranges of both the tame and aggressive foxes, while the scores of the backcross generation resegregate. These measures can thus be used for QTL mapping to explore the genetic basis of tame and aggressive behavior in foxes, which should provide new insights into the mechanisms of mammalian behavior and canine domestication.

Keywords

Canidae; Vulpes vulpes; interspecies tameness; attack; domestication

Introduction

The genetic basis of mammalian behavior has been studied in a limited set of species (Flint *et al.*, 2005; Kendler and Greenspan, 2006). The model choice was determined in large part by such criteria as availability of information on the genome, genetic tools, well defined genetically inherited behavioral phenotypes, and the opportunity for experimental breeding. Previously, among mammals, mainly rodent species could satisfy all these requirements. Information on the genomes of species representing different mammalian groups (http://www.broad.mit.edu/mammals/) has become available in the last few years. This progress in genome biology now creates an opportunity for genetic investigations in species

*Corresponding author: Anna V. Kukekova, James A. Baker Institute for Animal Health, Cornell University, Ithaca, NY 14853, Work: (607) 256−5693, Fax: (607) 256−5608, e-mail: avk5@cornell.edu.

that are well-established as models for behavioral studies but, until recently, lacked the major genetic tools (Scott and Fuller, 1965; Trut, 1999; Williamson *et al.*, 2003; Suomi, 2006).

In a long-term experiment at the Institute of Cytology and Genetics of the Russian Academy of Sciences (ICG), specific strains of silver fox with markedly different behavioral phenotypes have been selected (Trut, 1980a, 1980b; Trut 1999, 2001 and Trut *et al.,* 2004). Foxes bred for docility demonstrate a friendly response to humans similar to that of domestic dogs. In contrast, foxes from a strain selected for aggressive behavior are aggressive toward humans and difficult to handle. Inter-specific aggression "in defense of the subject's own bodily integrity" is classified as defensive aggression (Blanchard and Blanchard, 2005). These tame and aggressive fox strains have been bred separately for over 40 generations under strong selection for their respective phenotypes, but in a manner designed to deliberately minimize inbreeding. Inbreeding coefficients during selection remained in the range 0.02−0.07 (Trut, 1999, 2001; Trut et al., 2004), and this low level of inbreeding has been confirmed in recent analysis with microsatellite markers (Kukekova et al., 2004). The genetic nature of these fox behavioral phenotypes is well established (Trut, 1980a, 1980b, 1999, 2001). Because these genetically determined behavioral differences segregate in very large pedigrees of a single species, they offer an opportunity to map and identify the genes responsible. The evolutionary closeness of fox and dog (Wayne *et al.*, 1997) enables the exploitation of canine molecular genetic tools, particularly the 7.6× dog genome sequence (Lindblad-Toh et al., 2005), to facilitate construction of a fox meiotic linkage map and to undertake genetic mapping in foxes (Kukekova et al., 2004, 2005, 2007)

Measurement of behavior presents several challenges, particularly in attempting to map the underlying genetic loci. Because behavior is clearly phenotypically complex, but also has indisputable heritable aspects, it likely reflects a complex interaction of multiple genetic loci and environmental factors. To identify genes responsible for behavioral variation, it is first highly desirable to identify, among the wide spectrum of behavioral expression, those specific, independent, and presumably simpler aspects that can be measured objectively and quantitatively and that can be demonstrated to be inherited.

In the course of selection of foxes for behavior, two scoring systems for assignment of fox behavioral phenotypes have been used: one for the tame and another for the aggressive population (reviewed in Trut, 1980a, 1980b, 1999, 2001; Kukekova *et al.,* 2005). Although the original farm-fox population showed a continuous variation in behavior from "relatively less aggressive and fearful" to "extremely aggressive", very quickly the phenotypes in the selected tame and aggressive populations no longer overlapped. Foxes from the tame population were scored by ranking them based on a repertoire of tame behaviors which were either shown or not during interaction with an experimenter under the standardized conditions. Scores of tame foxes reflect the intensity of the fox's friendly response toward the experimenter: the tamest foxes are assigned scores of 3.5−4.0; the least tame score 0.5−1.0. Behavioral assessment in the tame population was further refined by evaluating a comprehensive set of measures for scoring behaviors contributing to "tameness" by principal-components analysis (Vasilieva and Trut, 1990). In contrast, the major criterion for measuring behavior in the aggressive population was the critical distance between the experimenter and the caged animal when the animal first demonstrates an aggressive reaction and intensity of the fox aggressive response (Trut, 1980a, 1980b, 1999, 2001; Kukekova *et al.*, 2005). Animals demonstrating the most aggressive response to humans are scored −4; those showing the least aggressive response score −0.5. The systems for measuring behavior in the tame and aggressive strains yield objective and reproducible behavioral assessment of individuals in both populations and were used to select animals exhibiting the most tame and the most aggressive behaviors for breeding the next generation. The continued improvement in scores with selection over multiple

generations (Trut 1999, 2001) is the best evidence for the reliability and utility of these scoring systems for measuring behavior in the tame and aggressive strains.

To study the genetic basis of fox behavioral phenotypes, three-generation experimental pedigrees have been established by breeding tame to aggressive founders to produce an F1 generation, and then backcrossing to the tame strain (Acland *et al.,* 2004; Kukekova *et al.*, 2005, 2007). Assignment of behavioral phenotypes in F1 clearly demonstrates that the traditional scoring systems established for selection of foxes for behavior has limited resolution for measuring behavior as a continuous variable in the cross-bred pedigrees. Broadly, F1 foxes exhibit a wide range of behaviors; substantial percent of foxes had low values on both the "tame" and "aggressive" scales (Trut, 1980a, 1980b; Kukekova *et al.*, 2005). Furthermore, behavioral patterns characteristic of the founder populations become fragmented or reshuffled in the cross-breed offspring. Thus, before attempting to map or identify genes underlying behavioral variations segregating in these fox strains, we needed a high-resolution, objective, quantitative system that defined behavior of foxes from both the tame and aggressive strains, and that enabled assignment of behavioral phenotypes in both founder and experimental populations.

In the current study we developed and tested a new system for assignment of fox behavioral phenotypes. To capture those fox behavioral components which had been selected for in the development of the founder populations, this new system is rooted in the traditional behavioral tests developed at ICG (Trut *et al.*, 2004; Vasilieva and Trut, 1990). The behavior of the foxes was evaluated as in the traditional methods, and videotaped. A comprehensive primary set of binary (present, absent or yes, no) objective observations was then developed for scoring the physical manifestations of fox behavior during the test from video records. Statistical analyses, including principal-components analysis (PCA), were used to dissect out the independent, resegregating traits underlying the phenotypic variation expressed in these multiply correlated observations. To validate this new system for measuring behavior we evaluated the concordance between the ICG behavioral assignment and this new system. Moreover, a useful system for measuring behavior in experimental cross-bred pedigrees has to distinguish between the behavior of foxes from the tame and aggressive strains as well as cover a range of values for the F1 and backcross-to-tame generations that are intermediate to the parental strains. Consistency of this new system was tested between independent data sets obtained by different observers and the reproducibility was tested among repeated tests.

The system described herein provides an essential tool for quantitative analysis of these fox behavioral phenotypes. Together with the newly available tools for genomic research, this now makes it feasible to map the loci and genes implicated in tame and aggressive fox behavior. Determination of the genetic basis for specific behavioral phenotypes in these foxes promises to yield broader insights into the genetics of complex behavior and its underlying molecular mechanisms.

Methods

1. Animals and animal maintenance

We used silver-fox populations developed and maintained at the experimental farm of the Institute of Cytology and Genetics (ICG), Novosibirsk, Russia. The five fox populations investigated included the three primary fox strains developed and maintained at ICG:

    i. the selectively bred tame strain;

    ii. the selectively bred aggressive strain;

iii. a population of farm-bred foxes unselected for behavior; and two additional populations produced for this study:

iv. an F1 generation developed by crossing tame males to aggressive females; and

v. a backcross population produced by reciprocally breeding F1 foxes back to the tame strain . All foxes we used were born in March-April of 2003, were raised under consistent farm conditions and had similar interactions with people (mostly limited to maintenance procedures). Pups were caged with their mothers until they were 1.5 months old. Subsequently, all littermates were housed together without their mother until 2.5 months of age, after which each pup was moved to its own cage.

2. Testing and recording of fox behavior

Behavior of foxes from the five populations (tame, aggressive, unselected, F1 and backcross) was evaluated using a version of the traditional test "the standard test" developed at ICG in the course of selective breeding for behavior (Vasilieva and Trut, 1990; http://cbsu.tc.cornell.edu/ccgr/behaviour/index.html). The test was designed to evaluate fox responses to humans in situations with different levels of interaction between experimenter and tested animal. Foxes were tested in their home cages by an observer, in five steps:

- Step **1** (Approach) observer approaches the fox's cage;

- Step **2** (Stand) observer stands calmly near the closed cage but does not deliberately try to attract the animal's attention;

- Step **3** (Door) observer opens the cage door, remains nearby but does not initiate any contact with the fox;

- Step **4** (Touch) observer attempts to touch the fox;

- Step **5** (Exit) observer closes the cage door, then stays calmly near the closed cage.

Each test step (except Approach) was 1 minute long.

Each fox was tested at 5.5−6 months old, at least twice, and a subset of foxes was tested three times. No more than one test was given to any individual animal on the same day. In most cases, the period between tests was 1 day. All tests were performed between 10:30 am and 5 pm, but no earlier than 30 minutes after feeding. An interval of least 30 minutes separated testing of animals in neighboring cages. All tests were videotaped with a digital camcorder (Canon ZR90) and subsequently copied onto DVDs to maintain a permanent record.

Behavior of foxes in the current study was assigned using both the traditional ICG scoring system and the new improved quantitative method from video records, as described below.

3. Scoring fox behavior from video records

A detailed analysis of fox behavior during the test period was undertaken from video records. The behavior of a subset of foxes (~20 animals from each population) was surveyed from videotapes and a comprehensive list of 311 recordable observations of fox behavior was identified (see Supplementary Table 1), each of which could be scored in a binary fashion (e.g. presence or absence). Scoring fox behavior from video records using a list of identified traits was performed manually using a DVD player or WinDVD software.

To evaluate the location of the fox in the cage, the space in each cage was partitioned into six zones. Zones 1−2 are located in the front of the cage (zone 2 is the closest to the experimenter), zones 5−6 are at the back of the cage, zones 3−4 are in the middle. To produce traits scoreable in a binary fashion ("yes" or "no") traits were designed to record whether a fox ever came into

a particular zone (for example, "Fox came to zones 1−2") or whether fox spent more than a certain period of time in particular zones (for example, "Fox spent >40s in zones 1−2"). Traits which were recorded in more than one test step (for example, "Wagging tail", "Staying at the front door", "Fox spent >40s in zones 5−6") were treated as independent traits.

4. Data sets

Behavior of a subset of foxes from all five populations was evaluated for either the comprehensive set of 311 traits, or a selected subset of 50 traits (see section **5**, below). Foxes of both genders were represented relatively equally in all data sets. Four data sets were generated as follows:

Data Set 1—Data set 1 represents behavioral scores for a set of 95 foxes, at 5.5−6 months age, scored from videotapes by observer 1 (author AVK), for all 311 traits. The 95 foxes included 20 animals each from the tame, aggressive, unselected and backcross populations, and 15 F1 animals. These 15 F1 foxes were all produced by breeding tame males to aggressive females, and they were the only appropriately aged F1 foxes available at that time (in 2003). For each of the other populations, twenty foxes were randomly chosen from a larger list of about 120 available animals per population, using the Microsoft Excel random-number function. When it was difficult to determine on the tape whether the trait was demonstrated or not, the trait was recorded as a missing value.

Data Set 2—Data set 2 represents a revised version of data set 1, scored by observer 1, for the reduced 50-trait set. To generate a data set without missing values for all 95 foxes, all missing scores were re-evaluated on the video again by observer 1. If it was still difficult to resolve the trait's presence or absence, the trait was re-scored as "not shown" by default. This was done because even a single missing data point would eliminate that animal's entire data set from subsequent analysis.

Data Sets 3A, 3B, and 3C—Data set 3 represents behavioral scores for the 50-trait set previously selected, for a second set of foxes (n=95), evaluated by three different observers. This set included 20 different randomly chosen foxes per population other than F1, and the same 15 F1 foxes used in set 1 (because there were no other F1 foxes yet available). The first test (at 5.5−6 months of age) was analyzed from video records by three trained observers (observer 1 was the same as before (Data set 3A), observer 2 was a high-school student (Data set 3B), and observer 3 was another scientist associated with the fox study (Data set 3C)). When it was not clear on the tape whether the trait was demonstrated or not, the trait was scored as "absent" by default (consistent with data set 2).

Data Set 4—Data set 4 represents behavioral scores for the 50-trait set, from 75 additional foxes (20 different animals from each of the tame, aggressive and backcross populations, and the same 15 F1 animals). Each fox was tested on three different days at 5.5−6 months of age, and video records were scored by observer 1.

5. Statistical analysis

5a. Trait evaluation—Data set 1 was used to identify the most informative traits, for each testing step, from the comprehensive list of 311 traits originally observed in the video survey. For practical reasons, we considered it desirable to reduce the set of traits to 50 per test step prior to undertaking principal-components analysis (PCA).

First, the simple frequency distribution of each trait was examined. Any dichotomous trait with (arbitrarily) <10 observations of the less-frequent category was removed from further consideration; this assured that the trait was indeed variable among the foxes. After these

deletions for the steps approach, stand, door, touch, and exit, 14 of 22, 33 of 77, 44 of 83, 48 of 72, and 36 of 57 variables, respectively, remained. The remaining traits (175 of the original 311, with no more than 50 traits per test step) were then examined using a separate PCA for each step, using the software program STATISTIX™ 8 Analytical Software, 2003). Then, each principal component (PC) which explained both >5% of the variance and more than (100%/n) of the variance (where n = number of traits; as per Afifi and Clark, 2004 pp: 379, 382) was evaluated by Kruskal-Wallis (rank-based 1-way ANOVA) to test whether it could distinguish between the five populations of foxes used in the PCA. If the PC was not significant in the multiple-comparisons tests of the Kruskal-Wallis procedure at P<=5% (2-sided), the PC was not retained for further analyses.

If the PC passed each of these criteria, then Spearman's rank correlations (rsps) were run between the individual-trait values and the PC for each fox. Traits yielding rsps with absolute values >= 50% were deemed 'important' contributors to at least one of the PCs, and were thus retained for the second stage of PCA. For approach, stand, door, touch, and exit, respectively, there were 5, 18, 18, 26, and 22 traits retained for the second stage.

In the second stage, two more PCAs were run: one included all the traits retained from scoring steps approach, stand, and door; the other included the traits retained from touch and exit. The same process of identification of traits to retain was used as before. In the final stage, one last PCA was run with traits that had been retained after the second stage.

5b. Reproducibility investigations—Reproducibility of traits was investigated in several ways. To evaluate disagreement between trait scores in different data sets and among observers, kappas were calculated; significance was interpreted as $P \leq 0.001$ (i.e., with a Bonferroni-type correction to maintain experiment-wise alpha of 0.05). Kappas < 0.40 were judged to indicate poor concordance, regardless of statistical significance; kappas ≥ 0.60 were judged to indicate substantial agreement (Dohoo *et al.*, 2003 p. 92).

6. Heritability of PC1

We estimated the heritability of PC1 across all 4 populations using variance component analysis on a 21 generation pedigree inclusive for all tame, aggressive F1 and backcross animals. The additive heritability was estimated using the "polygenic" function of SOLAR (Almasy and Blangero, 1998). Maternal effects were estimated using the "house" option of SOLAR with the ID of the mother. This adds an additional variance component to the model coded as 1 for all foxes which share the same mother and 0 for all others.

Results

1. Identification of informative observations of fox behavior in the standard test

From video records of foxes being tested using the ICG "standard test" an initial comprehensive set of 311 physical manifestations of fox behavior recording specific fox actions: location in the cage and time spent there, body postures, positions of particular parts of the body, and sounds (see Supplementary Table 1) was identified. All subjective assessments of fox actions (e.g. "Fox is afraid" or "Fox demonstrates submissive behavior") were avoided. Twenty-two recordable observations of fox behavior for test step approach, 77 traits for step stand, 83 for step door, 72 for step touch, and 57 traits for step exit were included in the primary 311-trait set.

2. Principal-component analysis of fox behavior with selected traits

The initial data set, representing scores by observer 1 for 311 traits from a set of 95 foxes (data set 1), was winnowed by statistical analysis to a reduced set of 50 traits (Table 1). "Test time"

and "gender" did not contribute importantly in any PC1. None of the traits from test step "approach" met the criterion of informativeness and this step was excluded from further analysis. For the remaining four steps (stand, door, touch, and exit) 9, 13, 14, and 14 traits were retained, respectively. PCA using the 50-trait set identified three PCs which each explained >5% of the total variance (experimental cutoff) and distinguished among populations (each of these PCs had P<0.023 in the Kruskal-Wallis tests). PC1 accounted for 48.4% of variance and distinguished the tame population from aggressive, unselected and F1; and also the backcross from unselected and aggressive.

2a. Comparison data sets with and without missing values—PCA with the same 50 traits and foxes as in data set 1 but with resolution of missing values (default = "absent") identified three PCs, each of which explained >5% of total variance and distinguished among fox populations (all P<=0.0041 in the Kruskal-Wallis tests). As in data set 1, PC1 from data set 2 explained a large portion of the recorded variation in fox behavior (46.4%) and made the same distinctions between populations (Figure 1). All traits contributing to PC1 using data set 2 loaded in the same direction as for PC1 using data set 1 (Table 1 and Supplementary Table 2) and were within 0.02 (absolute value) of the loading value in data set 1. The Spearman rank correlation test identified 48 traits judged to be "important" ($r_{sp} >= |0.5|$) for PC1 in data sets 1 and 2 (Table 1). Of the "important" traits, 47 were common to both datasets. All such identified "important" traits contribute significantly to PC1 (P=<0.0001).

3. Reproducibility of the PCA between data sets

PCA of behavioral scores on the mostly new set of foxes identified four PCs in data set 3A (observer 1; PC1 had P<0.0001, and the other three had P<0.08 on Kruskal-Wallis tests) and three PCs in data set 3B (observer 2; P values were <0.0001, 0.03, and 0.035 for PC1, PC2, and PC3 respectively). PC1 explained over 44% of behavioral variation in both data sets and distinguished among slightly different populations. PC1 in data set 3B distinguished the same populations as did data sets 1 and 2. PC1 in data set 3A differed from the other three only in that backcross and F1 were detected as different.

Animals with the highest PC1 values acted as if eager to establish human contact; foxes with the lowest PC1 values showed aggression towards humans. Trait signs for all 50 traits in PC1 were the same in data sets 1, 2, 3A, and 3B; 39 traits were judged to be important for PC1 in all four data sets (Table 1). We interpreted consistency and reproducibility of these results (i.e., which traits were judged to be important, and the contributory direction of those traits) as insensitivity of the scoring process to which individual foxes were scored (observer 1 between data sets) or to which observer did the scoring (observer 1 and 2 on the same data set). That is, the behavior measured by PC1 was consistently and reproducibly exhibited and scored, and differed primarily according to the population of origin.

4. Reproducibility of inter-observer judgments

For the 50 traits scored in data set 3, 41 showed substantial agreement (kappas >=0.60) between observers 1 and 2, and 34 traits between observers 1 and 3 (Table 1). Two traits (d17 and e56) had poor concordance (kappas <0.40) between observers 1 and 2, and four traits (s38, d17, e20, and e30) had poor concordance between observers 1 and 3 (Table 1). Both poor concordance and non-significance of the kappas suggest problems with inter-observer variation for those five traits.

5. Test repeatability

Test repeatability between days 1 and 2, and between days 1 and 3 showed acceptable or better concordance (kappas ≥ 0.40) for 42 and 32 traits, respectively (Table 1). Eight traits showed poor concordance (kappas <0.40) between tests given on days 1 and 2, and 18 traits showed

poor concordance between days 1 and 3 (Table 1). Trait d17 (sniffing floor/air) showed poor concordance both between observers (data set 3) and between tests (data set 4). Traits e20 (running in circle) and e56 (leaning on the right wall) had poor kappas in three of four tests. None of the 14 traits from step touch showed poor concordance in data sets 3 or 4.

6. Comparison of the ICG scoring system with the PC-based method

Concordance of the PC-based method with the traditional ICG scoring system was observed (Figure 2) by comparison of assigned behavioral phenotypes for 40 animals from each of the tame and aggressive populations (data sets 2 and 3A) using both approaches. Foxes from the tame population had higher PC1 values than foxes from the aggressive population. Animals with the highest PC1 values appeared eager to establish human contact while foxes with lowest PC1 values showed aggression towards humans. Thus, PC1 represents a quantitative representation of the more subjective traditional scoring system.

7. Segregation of PC1 in fox populations

PC1 demonstrated remarkable stability in trait composition among data sets. Most importantly, this PC clearly distinguishes foxes from the tame and aggressive populations, and shows resegregation in the backcross generation (Figure 1). Although the sample of F1 foxes in the current study was relatively small (15 F1 animals were included in the analysis) a wide distribution of F1 PC1 values was observed (Figure 1). PC1 values in the F1 generation cover a range of values intermediate between tame and aggressive animals. PC1 values in the backcross-to-tame foxes cover a range of values between the F1 and tame populations. The estimated heritability was 1 (p-value 1e-10) without a maternal effect and 1 (p-value 1e-7) with the maternal effect. The maternal effect was not significant. The heritability estimates in the individual populations were: tame (n=40) $h^2 = 0.14$ (p-value .36), aggressive (n=40) $h^2 = 0$ (p-value 0.5), backcross (n=39) $h^2 = 1$ (p-value 1e-7), F1 (n=15) was not estimated due to small sample size. The estimation of heritability in these populations is most likely biased by dominance and epistatic effects.

Discussion

The approach described herein precisely and robustly measures fox behavior as quantitative phenotypes -- principal components -- which capture the heritable differences in behavior among individual foxes and different fox populations. These continuous variables cover the entire range of behavioral variation characteristic of all the different populations, including the 3-generation experimental pedigrees. PC1, specifically, provides a valid and reproducible measure of fox behavior, suitable for quantitative genetic experiments including QTL mapping. PC1 stands in contrast to the standard ICG scoring system that was subjective and was based on two separate scoring systems and therefore, was indeterminate in the range of behaviors (between tame and aggressive) encountered in F1 animals derived from crosses between tame and aggressive parents.

From a preliminary set of 311 binarily scoreable behaviors, we defined the fifty most important traits, which were valid, reproducible and explained a substantial proportion (~50%) of variation in fox behavior. All 50 traits contributed significantly (P<0.0001) to PC1 and 39 traits were well correlated to PC1 in all datasets (Table 1). We acknowledge that five traits had poor reproducibility. Twenty-nine traits described clear aspects of fox behavior and could be scored reliably by observers after even a very short training period, whereas the other traits require additional training. Tame and aggressive traits were inversely correlated and defined a tame-aggressive axis that is remarkably stable in trait composition (loadings). In all data sets, PC1 distinguished the tame population from the F1, aggressive and unselected populations (figure 1).

As expected, the frequency of trait observations differs among populations. Almost every tame fox was wagging its tail (s37 and d21) during the test. In the contrast, traits associated with aggressive behavior (attack (t37), attack alert (t38), pinned ears (t40), aggressive sounds (t46) —see Supplementary table 1 for description) were never observed in the tame population in any data set. Eleven traits associated with tame behavior were never observed in the aggressive population. No F1 foxes exhibited behavior traits typical of the most tame animals from the tame population (s7, d22, d24, t27—see Supplementary table1 for description) but many do express other traits associated with tame behavior as well as some of the traits associated with aggressive response. The only trait which was not observed in the backcross-to-tame population in any data set was "attack" (t37).

Seventeen traits were not observed in the unselected population. Among these, eight (s7, d7, d12, t10, t13, t19, and t27 —see table1) were also not observed in the aggressive population and thus represent traits associated with a friendly fox response to humans. Other traits not observed in unselected foxes are those recording the location of the animal in the front part of the cage and/or a fox approaching an experimenter. Unselected foxes (more often than foxes from other fox populations) move back when the observer approaches the cage (s37), spend more time in the back part of the cage (s39), and move to the back part of the cage when the experimenter opens the cage door (d5). Compared to foxes from the aggressive population, unselected foxes show more sniffing of the front portion of the cage (d18), and more rear demonstrate attacking (t37) and/or expressing aggressive sounds (t46). We interpret these behaviors to be fear based. Similar observations of fearful behavior have been reported for farm-bred foxes in Finland (Harri *et al.*, 2003).

We tested the sensitivity of the trait to day-to-day variation within the fox. The high number of inconsistent traits observed during step stand might indicate that by the second test, foxes are familiar with the test structure and anticipate the human contact during steps door and touch. Analysis of larger data sets would be required to test whether animals which are eager to establish human contact begin to show tamer behavior during step stand with experience, while foxes that resist human contacts start to express more anxious behavior during step stand in the second and third test.

We hypothesize that the low consistency of step exit between days 1 and 3 is associated with animal experience. Some tame foxes might stop eager human contact when they learn that step exit is the last step of the test and they will not be petted any more. At the same time, some foxes from the aggressive population start to show less fearful/aggressive behavior when they learn that after the cage closes (step touch) people are not a danger anymore. This can also provide some information about fox trainability and memory. To test this hypothesize rigorously a larger data set is needed.

Step touch shows the most consistency among all test steps. This is the test step when contact between the fox and the experimenter is most intense and it shows that under these "extreme conditions" few changes are made with experience.

Factors that represent the tame-aggressive axis have not been reported in the literature for dogs, although there are important parallels to factors representing stranger-directed aggression, as described by Serpell and Hsu (Serpell and Hsu 2001; Hsu and Serpell, 2003) as well as factors representing sociability and aggressiveness described by Svartberg and Forkman (2002). In fact, this factor might explain (to a large extent) differences in the behavioral responses to humans that are observed between domesticated and wild species. The strains of tame and aggressive foxes described here provide an opportunity to understand the genetic differences underlying these behavioral phenotypes and should shed light on the genetic mechanisms involved in domestication of dogs and other species.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Acland, GM., et al. Resegregating behaviors in the Silver Fox.. A model system for mapping sociability. 2004. http://www.ashg.org/genetics/abstracts/abs04/f2300.htm

Afifi, A.; Clark, VA.; May, S. Computer-aided multivariate analysis. Chapman and Hall/CRC; 2004. p. 489

Almasy, L.; Blangero, J. Multipoint quantitative trait linkage analysis in general pedigrees.; Am J Hum Genet. 1998. p. 1198-1211.http://www.sfbr.org/solar/

Blanchard RJ, Blanchard CD. Some suggestions for revitalizing aggression research. Novartis Found Symp 2005;268:4–12. [PubMed: 16206872]

Dohoo, I.; Martin, W.; Stryn, H. Veterinary Epidemiologic Research. AVC Inc.; Charlottetown, PEI, Canada: 2003.

Flint J, et al. Strategies for mapping and cloning quantitative trait genes in rodents. Nat Rev Genet 2005;6 (4):271–86. [PubMed: 15803197]

Harri M, et al. Behavioural and physiological differences between silver foxes selected and not selected for domestic behaviour. Animal Welfare 2003;12:305–314.

Hsu Y, Serpell JA. Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. J Am Vet Med Assoc 2003;223(9):1293–300. [PubMed: 14621216]

Kendler KS, Greenspan RJ. The nature of genetic influences on behavior: lessons from "simpler" organisms. Am J Psychiatry 2006;163(10):1683–94. [PubMed: 17012675]

Kukekova AV, et al. A marker set for construction of a genetic map of the silver fox (Vulpes vulpes). Journal of Heredity 2004;95:185–194. [PubMed: 15220384]

Kukekova, AV., et al. The Genetics Of Domesticated Behavior In Canids: What Can Dogs And Silver Foxes Tell Us About Each Other? Chapter 21.. In: Ostrander, EA.; Giger, U.; Lindblad-Toh, K., editors. The Dog and Its Genome. Cold Spring Harbor Laboratory Press; Woodbury NY: 2005. p. 515-537.

Kukekova AV, et al. A meiotic linkage map of the silver fox, aligned and compared to the canine genome. Genome Research 2007;17(3):387–99. [PubMed: 17284676]

Lindblad-Toh K, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 2005;438(7069):803–19. [PubMed: 16341006]

Scott, JP.; Fuller, JL. Genetics and the social behavior of the dog. The University of Chicago Press; Chicago and London: 1965. p. 468

Serpell JA, Hsu Y. Development and validation of a novel method for evaluating behavior and temperament in guide dogs. Appl Anim Behav Sci 2001;72(4):347–364. [PubMed: 11348683]

Suomi SJ. Risk, resilience, and gene × environment interactions in rhesus monkeys. Ann N Y Acad Sci 2006;1094:52–62. [PubMed: 17347341]

Svartberg K, Forkman B. Personality traits in the domestic dogs (Canis familiaris). Applied Animal Behavior Sciences 2002;79:133–155.

Trut LN. The Genetics and Phenogenetics of Domestic Behaviour. Problems in General Genetics 1980a; 2:123–136.(Proceeding of the XIV International Congress of Genetics)book 2

Trut, LN. Doctoral (Biol.) Dissertation. Institute of Cytology and Genetics; Novosibirsk, Russia: 1980b. The Role of Behavior in Domestication-Associated Changes in Animals as Revealed with the Example of Silver Fox..

Trut LN. Early Canid domestication: The Farm Fox Experiment. American Scientist 1999;87:160–169.

Trut, LN. Book: *The Genetics of the Dog.*. CABI; 2001. Experimental Studies of Early Canid Domestication.; p. 15-43.

Trut, LN.; Pliusnina, IZ.; Os'kina, IN. Genetika (Russ.). 40. 2004. An experiment on fox domestication and debatable issues of evolution of the dog.; p. 794-807.

Vasilieva LL, Trut LN. The use of the method of principal components for phenogenetic analysis of the integral domestication trait. Genetika 1990;26(3):516–24. [PubMed: 2354785]

Wayne RK, et al. Molecular systematics of the Canidae. Syst Biol 1997;46:622–53. [PubMed: 11975336]

Williamson DE, et al. Heritability of fearful-anxious endophenotypes in infant rhesus macaques: a preliminary report. Biol Psychiatry 2003;53(4):284–91. [PubMed: 12586447]
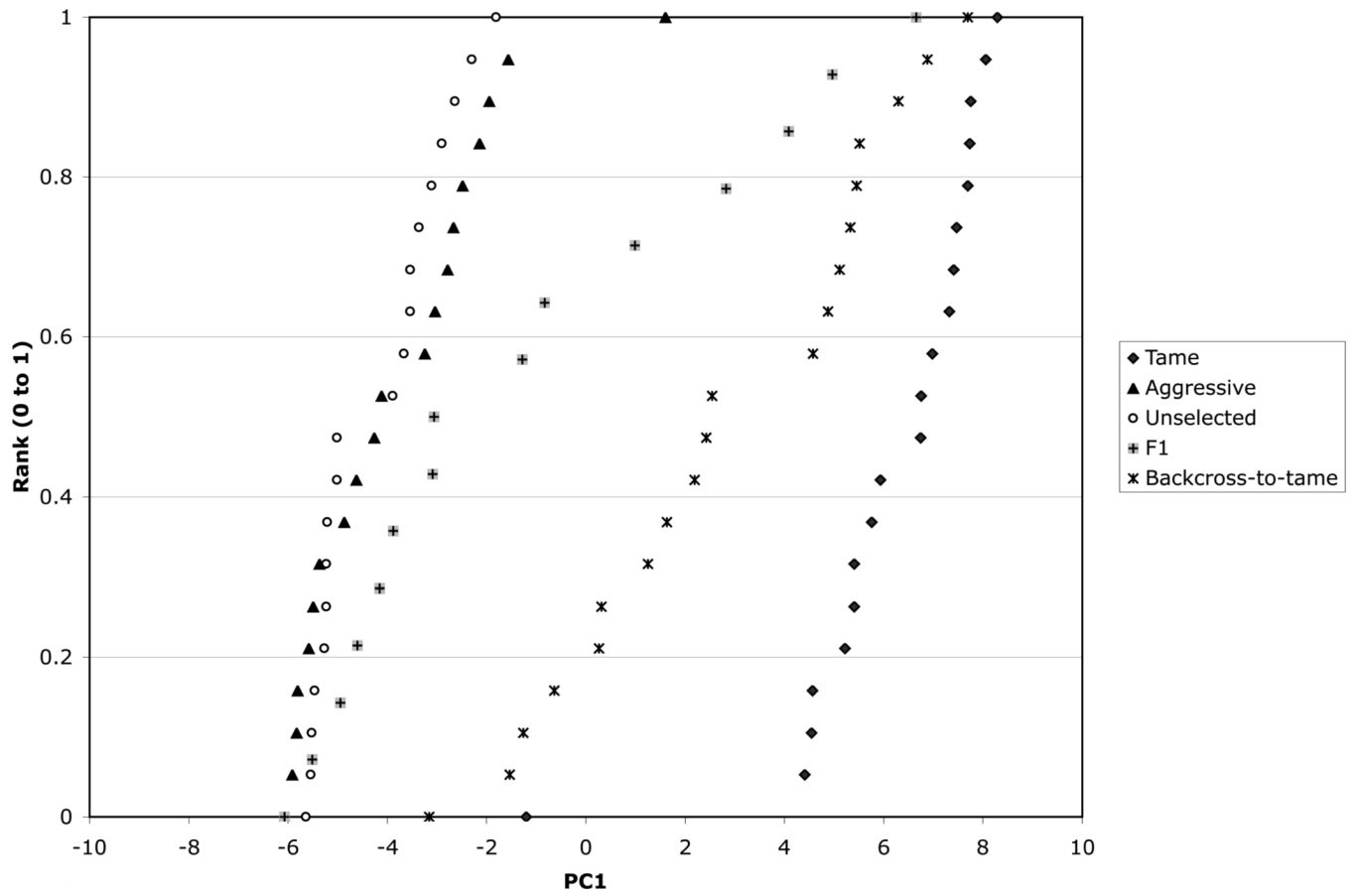
Figure 1.
Cumulative distribution of PC1 values among 20 tame, 20 aggressive, 20 unselected, 15 F1, and 20 backcross-to-tame foxes.
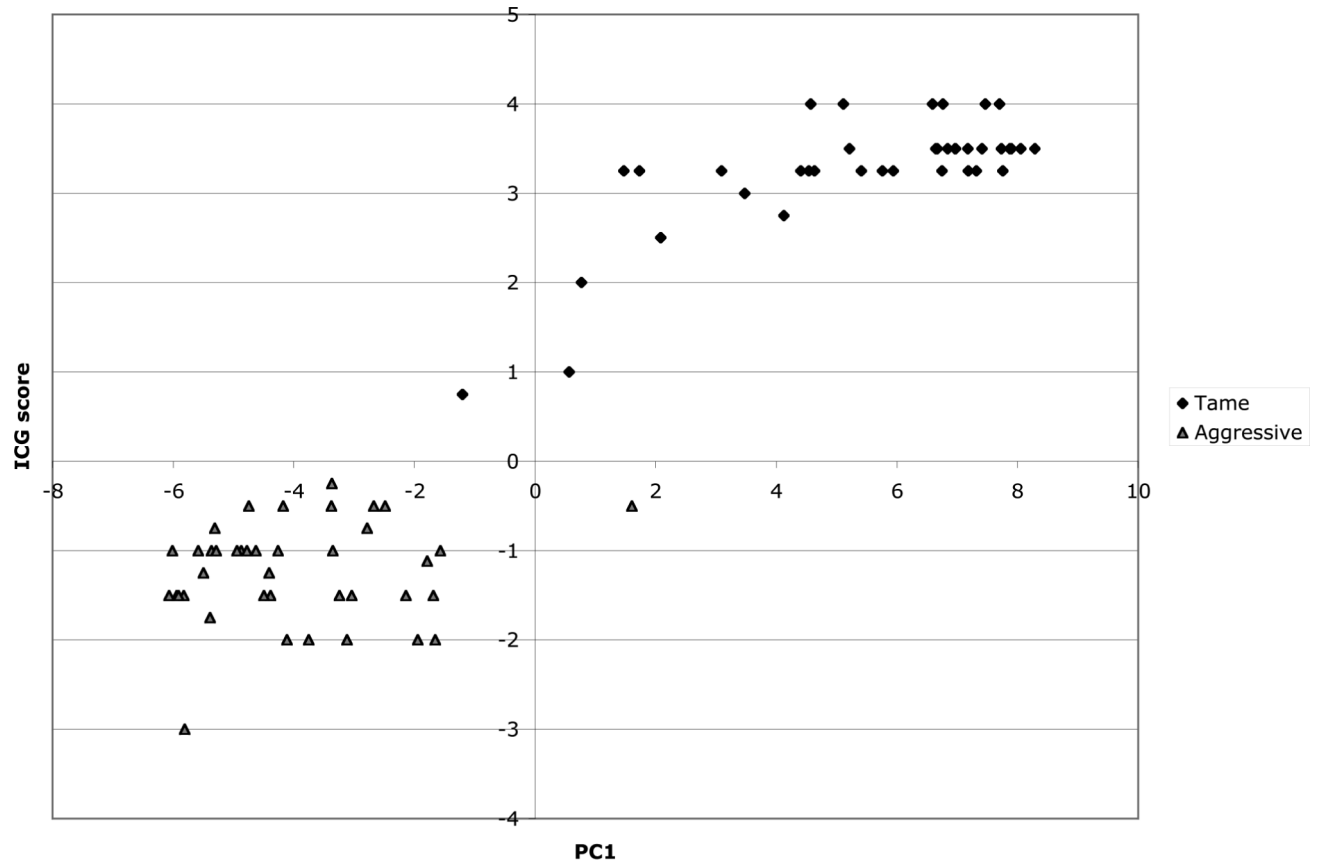
Figure 2.
Comparison of behavioral assignment using both the standard ICG scoring system and PC1, in 40 foxes from each of the tame (diamonds) and aggressive (triangles) populations.

Table 1

Analysis of 50 traits selected for behavioral assignment.

	Trait	Code	Data Set 1	Data Set 2	Data Set 3A	Data Set 3B	PC1 important traits	Kappas among observers		Kappas among tests (Data set 4)	
								Data Set 3A vs 3B	Data Set 3A vs 3C	Test 1 vs. Test 2	Test 1 vs. Test 3
1	Wagging tail	S7	+	+	(+)	(+)	2+	0.87	0.87	0.72	0.62
2	Touching cage door with nose	S12	+	+	+	+	4+	0.63	0.67	0.5	0.39
3	Sniffing the front door of the cage	S13	+	+	+	+	4+	0.86	0.71	0.35 ns	0.38
4	Staying at the front door of the cage	S16	+	+	+	+	4+	0.93	0.93	0.36	0.31 ns
5	Sitting in zone 2 looking at observer	S20	+	+	+	+	4+	0.82	0.72	0.26 ns	0.47
6	Moving back at least one zone during first 15 seconds (including step 1)	S37	(−)	(−)	−	−	2−	0.65	0.46	0.29 ns	0.32 ns
7	Spends at least 40 seconds in zones 1–2–3–4	S38	+	(+)	+	+	3+	0.73	0.34	0.49	0.36
8	Spends at least 40 seconds in zones 3–4–5–6	S39	−	(−)	−	−	3−	0.75	0.54	0.52	0.31 ns
9	Comes into zones 1–2	S49	+	+	+	+	4+	0.84	0.69	0.37	0.43
10	Fox moved immediately to zone 5 or 3–5	D5	−	−	−	(−)	3−	0.62	0.62	0.43	0.58
11	Fox approaches the hand for at least 40 seconds	D7	+	+	+	+	4+	0.73	0.75	0.62	0.41
12	Fox tries to nip the hand or pokes it with nose	D12	+	+	+	+	4+	0.92	0.87	0.68	0.53
13	Sniffing floor/air	D17	+	−	(−)	−	2−	0.19 ns	0.36	0.21 ns	0.18 ns
14	Sniffing the front wall/door	D18	+	+	+	+	4+	0.75	0.77	0.59	0.33 ns
15	Wagging tail	D21	+	+	+	+	4+	0.86	0.93	0.81	0.74
16	Ears horizontal/down for at least 10 seconds	D22	+	+	(+)	+	3+	0.69	0.89	0.8	0.88
17	Body shaking	D24	+	+	(+)	+	3+	0.51	0.64	0.61	0.62
18	Not on the floor of the zone 2 at all	D27	−	−	−	−	4−	0.67	0.86	0.68	0.6
19	Comes into zones 1–2	D30	+	+	+	+	4+	0.79	0.59	0.6	0.41
20	Comes to the hand and sniffing	D32	+	+	+	+	4+	0.73	0.75	0.62	0.57
21	Spends at least 40 seconds in zones 1–2–3–4	D39	+	+	+	+	4+	0.56	0.65	0.64	0.51
22	Spends at least 40 seconds in zones 5–6	D41	−	−	−	−	4−	0.66	0.61	0.54	0.42
23	Lying down during contact	T9	+	+	+	+	4+	0.77	0.55	0.68	0.57
24	Rolling on side or back during contact	T10	+	+	+	+	4+	0.73	0.49	0.4	0.6
25	Ears held horizontal/down	T13	+	+	+	+	4+	0.61	0.51	0.76	0.68
26	Fox allows the back of its neck to be touched	T14	+	+	(+)	+	3+	0.51	0.51	0.72	0.84
27	Fox allows its back to be touched	T15	+	+	+	+	4+	0.86	0.83	0.9	0.78
28	Fox allows its nose to be touched	T16	+	+	+	+	4+	0.94	0.98	0.81	0.86
29	Fox allows its head to be touched	T17	+	+	+	+	4+	0.87	0.93	0.91	0.85
30	Fox tries to hold the observer's hand in its mouth	T19	+	(+)	+	+	3+	0.75	0.51	0.59	0.47
31	Breathing loudly	T27	+	+	+	+	4+	0.64	0.73	0.65	0.5
32	Attack	T37	−	−	(−)	(−)	2−	0.73	0.43	0.71	0.66
33	Attack alert	T38	−	−	−	−	4−	0.56	0.71	0.71	0.72
34	Pinned ears	T40	−	−	−	−	4−	0.74	0.66	0.74	0.74
35	Aggressive sounds	T46	−	−	−	−	4−	0.7	0.87	0.77	0.67
36	Fox moved to zone 2	T49	+	+	+	+	4+	0.75	0.7	0.78	0.57

	Code	Trait	Data Set 1	Data Set 2	Data Set 3A	Data Set 3B	PC1 important traits	Kappas among observers		Kappas among tests (Data set 4)	
								Data Set 3A vs 3B	Data Set 3A vs 3C	Test 1 vs. Test 2	Test 1 vs. Test 3
37	E4	Fox in zone 2 during the first 5 seconds	+	+	+	+	4+	0.94	0.91	0.59	0.59
38	E5	Spends at least 30 seconds in zones 1–2	+	+	+	+	4+	0.77	0.49	0.55	0.45
39	E7	Staying at the front door	+	+	+	+	4+	0.96	0.98	0.46	0.32 ns
40	E9	Touching the front wall with fore feet	+	+	+	+	4+	0.66	0.83	0.52	0.27 ns
41	E13	Touching the door with nose	+	+	+	+	4+	0.67	0.73	0.6	0.56
42	E20	Running in the cage in a circle	(+)	+	+	(+)	2+	0.53	0.21 ns	0.39	0.31 ns
43	E26	Sitting in zone 2 looking at observer	+	+	+	+	4+	1	0.77	0.57	0.38
44	E30	Spends more than 40 seconds in zones 5–6	–	–	–	–	4–	0.79	0.37	0.51	0.32 ns
45	E32	Spends more than 40 seconds in zones 1–2–3–4	+	+	+	+	4+	0.58	0.46	0.59	0.45
46	E33	Initially spends more than 10 seconds in zones 5–6	–	–	(–)	–	3–	0.52	0.66	0.44	0.2 ns
47	E43	Comes into zones 1–2	+	+	+	+	4+	0.82	0.59	0.53	0.39
48	E49	Changed position in cage 5 or more times	+	+	+	+	4+	0.69	0.85	0.54	0.25 ns
49	E54	Did not come to the floor of zone 2	–	–	–	–	4–	0.8	0.72	0.54	0.37
50	E56	Leaning on right wall in zone 2	+	+	+	(+)	3+	0.23	0.64	0.39	0.24 ns

Traits judged to be important and contributing to PC1 with a positive sign are marked by "+"; with negative sign are marked by "-". Signs of traits that did not meet the criterion for importance are marked as (+) or (−). Traits judged to be important for PC1 in all data sets with a positive sign are marked as 4+, in three data sets as 3+, etc.; traits which are important for PC1 in all data sets with a negative sign are marked as 4−, in three data sets as 3−, etc.