Лисицы и их геном
Why some animals are so nice?

- We still know little about molecular mechanisms influencing animal responses to humans.
- We can easily establish close interactions with some animals but not with others.
- Can animals which are more kin to interact with humans help us to learn about mechanisms of social behavior?
What is domestication?

- Domestication is distinct from mere taming of wild-born animals.
- It is the process whereby an animal is transformed from a life in a wild to a life under some control of humans.
- It involves animal transformations through many generations.
**Time scale of domestication**

- **Dog** (*Canis familiaris*): 15,000 BP or earlier
- **Sheep** (*Ovis aries*): ~11,000 BP
- **Pig** (*Sus scrofa*): ~9,000 BP
- **Cat** (*Felis silvestris*): ~7,500 BP in Cyprus, ~3,000 BP in China
- **Horse** (*Equus ferus caballus*): ~4,000-5,000 BP
- **Cattle** (*Bos taurus*): ~8,000 BP
- **Humped Cattle** (*Bos indicus*): ~5,000 BP
Physiological and morphological changes

“By a domesticate, I mean a species bred in captivity and thereby modified from its wild ancestors in ways making it more useful to humans who control its reproduction and food supply.”

Jared Diamond (Guns, germs, and steel)
Behavioral changes associated with domestication

- Reduced aggressiveness
- Reduced sensitivity to environmental changes
- Increased social tolerance among conspecifics

Courtesy of Darya Shepeleva
Genetics of domestication

- New sequencing technologies allow comparison of genomes and gene expression patterns of domesticated species and their wild counterparts.

- Finding genes and gene networks influencing domesticated behavior will inform us about physiology and neurobiology of animal social behavior.
Red fox (*Vulpes vulpes*) is an experimental model of animal domestication

The Institute of Cytology and Genetics of the Russian Academy of Sciences

1959

Selection for tame behavior

Dr. Dmitry Belyaev

Dr. Lyudmila Trut
Red fox (*Vulpes vulpes*) is an experimental model of animal domestication.

The Institute of Cytology and Genetics of the Russian Academy of Sciences

1959

Dr. Dmitry Belyaev  
Dr. Lyudmila Trut

Selection for tame behavior
Farm-fox experiment

Farm Foxes

1959

1970
Step A: STAY
Step B: OPEN DOOR
Step C: TOUCH
Step D: STAY
Advantages of the Fox Model

- Selected solely for behavior
- Genetic inheritance of the tame and aggressive phenotypes was carefully confirmed.
- The populations are outbred and provide an opportunity for high resolution genetic mapping.
- All foxes live under standard farm conditions and their behavior can be tested using standardized methods.
The red fox genome was sequenced at 93X and assembled using SOAPdenovo2

- Total length: 2,495,544,672 bp
- Scaffold N50: 11,799,617 bp
- Number of scaffolds: 676,878
- 94% of the genome is in top 500 scaffolds ranging from 47.7 Kb to 55.7 Mb

Kukekova et al., Nature E&E, 2018
Construction of Red Fox Chromosomal Fragments

Fox-Dog Syntenic Blocks

Cat-Dog Syntenic Blocks

RACA

128 RACA Fragments

Comparing the RACA fragments to the fox meiotic map using mapped markers

40 fox chromosomal fragments

Rando et al., 2018
Comparative analysis of fox populations

- Re-sequenced 10 foxes from each of the three populations maintained at the Institute of Cytology and Genetics in Novosibirsk:
  - Tame
  - Aggressive
  - Conventional

- Identified 8,324,814 fox SNPs
Fox population structure

(a) PCA plot showing the distribution of Tame, Aggressive, and Conventional foxes across PC1 and PC2.

(b) Circular network representation of individual foxes, with connections indicating relationships.

(c) Heat maps for different values of K (2, 3, 4, 5), illustrating the clustering of foxes into Aggressive, Conventional, and Tame categories.
Search for targets of selection for behavior

- Pooled heterozygosity ($Hp$) analysis: identification of genomic regions with reduced heterozygosity in each of the three populations.

- $F_{ST}$ analysis: identification of genomic regions with extreme divergence between the populations.
**$F_{ST}$ and pooled heterozygosity ($Hp$)**

- Window based analyses
  - Window size: 500 Kb
  - Average number of SNPs per window: 1784
- $Hp$ cut off in each population was established by permutations (Qanbari et al., 2012) and simulations
- $F_{ST}$ cut off ($>0.458$) was established using gene drop simulations
**$F_{ST}$ and $Hp$ analyses in tame and aggressive populations**

- Number of windows with $Hp$ at $p < 0.0001$
  - Tame ($HpT$): 96 (combined: 30)
  - Aggressive ($HpA$): 60 (combined: 19)

- Number of windows with $F_{ST} > 0.458$
  - 276 (combined: 57)
$F_{ST}$ and $Hp$ analyses in tame and aggressive populations
Comparison of behavioral QTL intervals with genomic regions identified in $F_{ST}$ and $Hp$ analyses
Behavioral trait D.PC1

Step A: STAY
Step B: OPEN DOOR
Step C: TOUCH
Step D: STAY
Behavioral trait D.PC1

<table>
<thead>
<tr>
<th>PC</th>
<th>Behavioral category</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>D.PC1</td>
<td>Avoiding the front part of the cage</td>
<td>50% of behavioral variance at step D</td>
</tr>
<tr>
<td></td>
<td>Aggression</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neutral behavior</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Located or moving to the front part of the cage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exploratory behavior</td>
<td></td>
</tr>
</tbody>
</table>
Fine mapping

- Fox chromosome 15 contains a single window with $F_{ST} > 0.458$ which is located within quantitative trait loci (QTL) interval for the trait D.PC1.

- The same window has low $Hp$ in both tame (0.20) and aggressive (0.23) populations.
Fine mapping

- Genotyping 25 small insertions/deletions located in a 5 Mb interval in tame (64) and aggressive (70) foxes.
Central (SorCS1) region

- Haplotype analysis identified three most common haplotypes:

- “Tame (olv)” haplotype was observed only in tame population (61%).

- “Aggressive-1 (trq)” – 47% in aggressive population and 7% in tame population;

- “Aggressive-2 (lav)” – 37% in aggressive population and 2% in tame population.
Distribution of olv, trq, and lav haplotypes in fox populations
Effect of \textit{olv}, \textit{trq}, and \textit{lav} haplotypes on behavior in F2

- F2 population demonstrates a wide spectrum of behavior
- 25 small insertions/deletions were genotyped in F2 pedigrees previously used for QTL mapping (536 F2 foxes).
- D.PC1 scores were compared for F2’s homozygous for \textit{olv}, \textit{trq}, and \textit{lav} haplotypes.
SorCS1 haplotypes and behavior of F2s

- Mean D.PC1 values:
  - Homozygous olv: 0.068
  - Homozygous lav: -0.546

- Significant differences among three groups:
  - Kruskal-Wallis, p=0.03

- A post-hoc Dunn’s test with Benjami-Hochberg correction achieved a p=0.014 for the comparison of olv and lav homozygotes.
SorCS1 is a novel gene for social behavior

- The *olv* haplotype is observed with frequency ~60% in tame population and not found in aggressive and conventional populations.
- The *olv* haplotype influences behavior of F2 foxes in an expected direction.
- *SorCS1* is a strong positional candidate gene for behavioral QTL on fox chromosome 15.
SorCS1

- SorCS1 (sortilin-related receptor CNS expressed 1) is a member of the Vsp10p-D (vacuolar protein sorting/targeting protein 10-domain) type I transmembrane receptor family, which includes five genes and plays an important role in intracellular trafficking.

- Mutations in SorCS1 have been found to be associated with late Alzheimer disease, diabetes, and autism.

- SorCS1 is a key regulator of synaptic receptor trafficking. Neurexins and AMPA glutamate receptor are the major proteins sorted by SorCS1 (Savas et al., 2015).
Genes in fox regions – Gene Ontology

- 971 genes identified in all significant windows:
  - **HpT**: damaged DNA binding, single guanine insertion binding, guanine/thymine mispair binding
  - **HpA**: cytokine activity, B cell mediated immunity, carbohydrate binding, receptor binding
  - **F_{ST}TA**: natural killer cell activation, B cell mediated immunity, interleukin-1 receptor binding; carbohydrate binding, clathrin-coated vesicle
### Genes associated with human behavioral disorders found in fox regions

<table>
<thead>
<tr>
<th>Genes from four SFARI categories</th>
<th>SFARI category</th>
<th>Bipolar disorder associated genes</th>
<th>Genes known to be involved in mouse aggression</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKAP9</td>
<td>suggestive evidence</td>
<td>BAZ2B</td>
<td>CNGA2</td>
</tr>
<tr>
<td>AMPD1</td>
<td>suggestive evidence</td>
<td>CACNA1C</td>
<td>DCT</td>
</tr>
<tr>
<td>APH1A</td>
<td>suggestive evidence</td>
<td>CHRNA7</td>
<td>KCNJ3</td>
</tr>
<tr>
<td>ATP10A</td>
<td>suggestive evidence</td>
<td>GNG4</td>
<td>NCAM1</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>syndromic</td>
<td>GPR50</td>
<td>PAK7</td>
</tr>
<tr>
<td>CHRNA7</td>
<td>suggestive evidence</td>
<td>IQGAP2</td>
<td>PRNP</td>
</tr>
<tr>
<td>CNTNAP2</td>
<td>syndromic, strong candidate</td>
<td>NCAM1</td>
<td>TACR1</td>
</tr>
<tr>
<td>GRIN2B</td>
<td>high confidence</td>
<td>NTF3</td>
<td></td>
</tr>
<tr>
<td>KAT2B</td>
<td>strong candidate</td>
<td>PTPRO</td>
<td></td>
</tr>
<tr>
<td>MAGEL2</td>
<td>syndromic, strong candidate</td>
<td>RASGRF2</td>
<td></td>
</tr>
<tr>
<td>MYO9B</td>
<td>suggestive evidence</td>
<td>RBFOX1</td>
<td></td>
</tr>
<tr>
<td>PIK3R2</td>
<td>syndromic</td>
<td>SCAMP1</td>
<td></td>
</tr>
<tr>
<td>PLCB1</td>
<td>suggestive evidence</td>
<td>ZNF385D</td>
<td></td>
</tr>
<tr>
<td>RBFOX1</td>
<td>suggestive evidence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Reduced heterozygosity in aggressive foxes in the region syntenic to the Williams-Beuren syndrome in humans

<table>
<thead>
<tr>
<th>HpT</th>
<th>HpA</th>
<th>HpC</th>
<th>Fox genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.266</td>
<td>0.249</td>
<td>0.241</td>
<td>FBXO24, PColCE, MOSP3, POP7, EPO, EPHB4, SLC12A9, TRIP6, SRRT, UFS1, MUC3A, RPL7A, TRIM56, VGF, MOGAT3, PLOD3, CLDN15, TFR2, GIGYF1, ACHE, SERPINE1, AP1S1, NAT16, ZNHIT1, FIS1, ACTL6B</td>
</tr>
<tr>
<td>0.282</td>
<td>0.239</td>
<td>0.229</td>
<td>RPL7A, TRIM56, VGF, MOGAT3, PLOD3, CLDN15, RABL5, SERPINE1, AP1S1, NAT16, ZNHIT1, FIS1, COL26A1</td>
</tr>
<tr>
<td>0.322</td>
<td>0.087</td>
<td>0.226</td>
<td>RABL5, MYL10, COL26A1, CUX1</td>
</tr>
<tr>
<td>0.289</td>
<td>0.06</td>
<td>0.253</td>
<td>MYL10, CUX1</td>
</tr>
<tr>
<td>0.267</td>
<td>0.06</td>
<td>0.295</td>
<td>SH2B2, ORAI2, ALKBH4, LRWD1, POLR2J, POLR2J3, UPK3B, RASA4, CUX1, DTX2, PRKRP1</td>
</tr>
<tr>
<td>0.281</td>
<td>0.051</td>
<td>0.327</td>
<td>SH2B2, ORAI2, ALKBH4, LRWD1, POLR2J, POLR2J3, UPK3B, HSPB1, RASA4, ZP3, SRCRB4D, YWHAG, SRRM3, CUX1, DTX2, STYXL1, PRKRP1</td>
</tr>
<tr>
<td>0.342</td>
<td>0.05</td>
<td>0.342</td>
<td>HSPB1, CCL24, CCL26, ZP3, SRCRB4D, YWHAG, SRRM3, POR, STYXL1, RHBD2H, HIP1</td>
</tr>
<tr>
<td>0.348</td>
<td>0.094</td>
<td>0.341</td>
<td>CCL24, CCL26, POR, STYXL1, RHBD2H, HIP1, POM121, NSUN5, TRIM50, FKB6</td>
</tr>
<tr>
<td>0.304</td>
<td>0.197</td>
<td>0.33</td>
<td>HIP1, POM121, NSUN5, TRIM50, FKB6, FZD9, TBL2, BAZ1B, BCL7B, MLXIPL, VPS37D</td>
</tr>
<tr>
<td>0.285</td>
<td>0.271</td>
<td>0.317</td>
<td>FZD9, TBL2, DNAJC30, STX1A, ABHD11, CLDN3, CLDN4, WBSCR28, BAZ1B, BCL7B, MLXIPL, VPS37D, WBSCR22, WBSCR27</td>
</tr>
</tbody>
</table>
What does the fox say?

- Even a relatively simple behavioral pattern can be comprised by several behaviors with relatively independent genetic inheritance.

- Application of a combination of approaches (population and pedigree analyses) is powerful for finding genes influencing behavior.

- Genetics/genomics of domesticated behavior can provide a novel insight into molecular mechanisms involved in regulation of social behaviors.
Acknowledgements

- **University of Illinois**
  Jennifer Johnson
  Halie Rando
  Jessica Hekman
  Estelle Bastoues

- **BGI, China**
  Guojie Zhang
  Shiping Liu
  Xueyan Xiang
  Shaohong Feng
  Zijun Xiong

- **Royal Veterinary College, London, UK**
  Denis Larkin
  Marta Farre

- **ICG, Novosibirsk, Russia**
  Lyudmila Trut
  Anastasiya Kharlamova
  Irina Pivovarova
  Tatyana Semenova
  Anastasiya Vladimirova
  Darya Shepeleva
  Rimma Gulevich
  Yury Herbeck

- **IMCB, Novosibirsk, Russia**
  Violetta Beklemicheva
  Polina Perelman
  Natalya Serdyukova
  Aleksandr Graphodatsky

- **St. Petersburg State University**
  Stephen O’Brien
  Klaus Koepfli

- **Support**
  NIH GM120782
  ICiG
  Russian Fund for Basic Research
  UIUC Research Board
Felid & Canid Genome Assemblies

Assembly First Published: 2005
Technology: Sanger Sequencing
Chromosomes: Assembled
Version Used: CanFam3.1

Assembly Published: 2007
Technology: Sanger Sequencing
Chromosomes: Assembled
Version Used: FelCat5

Assembly Published: 2018
Technology: Short-Read Next-Generation Sequencing
Chromosomes: Not assembled.
676,878 scaffolds
Version Used: vv2.2

50-60 MYA
9-10 MYA
CARNIVORA
CARNITAE
Canis lupus familiaris
Vulpes vulpes
Felis catus
Reference-Assisted Chromosome Assembly (RACA) looks for conservation among species.

RACA

Fox-Dog-Syntenic Blocks

537 Conserved Fragments

Cat-Dog-Syntenic Blocks

Fox-Dog-Cat Phylogeny
RACA also integrates information from sequencing libraries.

Each RACA fragment has a known position in the dog chromosomes and fox scaffolds.
Comparing the RACA fragments to the fox meiotic map

The microsatellite markers used in the fox meiotic linkage Map have known positions on the fox chromosomes.
Comparing the RACA fragments to the fox meiotic map

We identified the positions of the markers on the RACA fragments. The markers allowed us to assign each RACA fragment a position along a fox chromosome.
Comparing the RACA fragments to the fox meiotic map
Comparing the RACA fragments to the fox meiotic map

If no markers mapped to the fragment, its position was determined based on fox-dog synteny.
Comparing the RACA fragments to the fox meiotic map

The size of the gaps between the RACA fragments was determined based on synteny with the dog.
Fox Chromosomes Aligned to Dog and Cat

The end result is 40 fox chromosomal fragments with corresponding syntenic positions in dog and cat.
Fox chromosomal fragment assembly

• 40 chromosomal fragments
• Deeply anchored to the dog genome...
Deep sequencing data

- 41 tame and 42 aggressive foxes were sequenced with >15x coverage per individual.
- Reads aligned against the fox genome with Bowtie2.
- SNVs called with GATK 4.
- 15,907,200 SNVs remained after filtering.
- One missense mutation identified in SorCS1 gene.
Association mapping of tame and aggressive foxes using k-mers (following Rahman et al. eLife 2018;7:e32920)

- 2,470,983,266 k-mers (21 bp long) were identified in tame and aggressive reads using Jellyfish.
- 9.23% of the k-mers were enriched in tame and 8.05% of the k-mers were enriched in aggressive samples.
- K-mers enriched in tame and aggressive samples were assembled separately using ABySS.
- K-mer assemblies were mapped against the dog genome using Bowtie2.
Tame enriched k-mers upstream SorCS1

• A region with multiple tame-enriched k-mers located approximately 55 Kb upstream from the SorCS1 gene in the dog genome was identified.

• The region is ~2kb in length (~CFA28: 19,171,000-19,173,000) and present only in tame foxes caring olv haplotype.
Olv haplotype upstream SorCS1 is different
Region upstream SorCS1 identified by k-mer mapping

~2 Kb insertion/deletion upstream of SorCS1 gene differentiates *olv* from other fox haplotypes in this region.