iDog

Michael Armas and Jonar Cowan
Department of Biology
BIOL 375

What is iDog?

- Integrated research resource for domestic dogs and wild canids
- Based out of China
- Several data types:
  - Genes
  - Genomes
  - SNPs
  - Breed/Disease Traits
  - Gene Expression

What is iDog?

- Links to other specific databases
  - Connected to NGDC (which links to over 40 databases)
- Public release on January 17th, 2017
- Updated every few months
  - Last updated June 21st, 2019

Database Specifics

- Phenotypic data gathered from Kennel Clubs/Unions
- Curated from dog research communities and public resources
  - CIDD
  - OIMA
  - TheDogPlace
  - Ensembl
  - UniProt
- Maintained by National Genomics Data Center
- Funded by various associations and programs of the Chinese Academy of Science

Contents

- Over 60 million data points
- Provides literature about specific species of canids
- Provides tools for genomic data visualization and analysis

Data contents	Data statistics
Gene	
Gene integrated from Ensembl database	32 220
Gene associated with Dog disease	229
SNP	
Individuals	127
Non-redundant SNPs identified	42 871 184
Non-redundant SNPs annotated in gene	22 031 720
Phenotype & disease	
Standard Breeds	473
Diseases	783
Genotype–phenotype pairs (G2Ps)	594
Genome	
Scaffolds for wolf	581
Scaffolds for Dhole	749
Gene expression	
RNA-Seq projects	7
Experiments	83
Tissues	5
Genes well annotated	27 534
Gene ontology (GO)	
Molecular function	
Genes	14 361
Annotations	60 030
Biological process	
Genes	15 079
Annotations	103 120
Cellular component	
Genes	15 884
Annotations	65 214
Literature	
Papers and books	6 535

Utility for the Public

- Targeted Audience has background in biology and genetics
- Some information is understandable by those without experience
- Phenotype database easily understandable
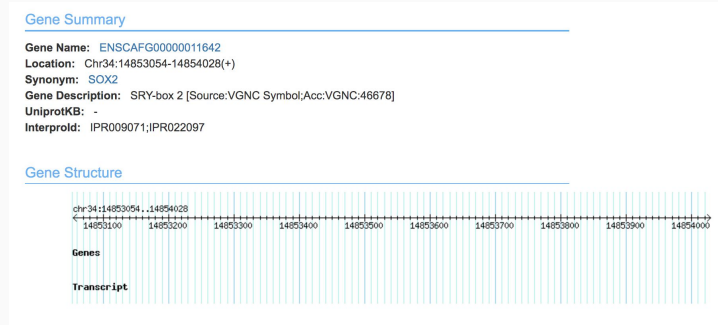
User Friendliness

- The layout design is organized, but lacks some genetic information compared to other databases.
- The database definitions can be found when searching up genes to provide basic information
- Navigating this site could be difficult if user does not understand biological terminology
- Tutorials are not helpful

User Friendliness (cont.)

- The search function only works to find gene symbol, gene name, and gene location
- Example searches are provided for quick understanding
- However, a general search bar is not found and requires you to look for tutorials/help

Welcome to iDog !

iDog, an integrated resource for domestic dog (*Canis lupus familiaris*) and wild canids, provides the worldwide dog research community a variety of data services. This includes Genes, Genomes, SNPs, Breed/Disease Traits, Gene Expressions, GO Function Annotations, Dog-Human Homolog Diseases and Literatures. In addition, iDog provides Online tools for performing genomic data visualization and analyses.

Search: Gene Symbol ▼ | e.g. ENPP1 | 🔍

e.g. MERTK, ENSCAFG00000000338, Chr1:14702150-37601000(+)

FAQ/Tutorials

- Finding the help section is difficult and offers a limited amount of information
- Tutorials are too little info, FAQ is too much

Tutorial

1. Gene Query

Users can use filter restricting their gene queries with our gene query form. Detailed information of the results in iDog are presented in 10 categories, including Gene Summary, Gene Structure, SNP Information, Exon Information in Dhole Genome Assembly, GWAS Catalog, Dog Disease Information, Expression Information, Homolog Gene, Gene Ontology and Genome Browser.

Click here for a detailed tutorial.

2. Dog Genome Database

Dog Genome Database (DogGD) is a data container for the genome assembly information of wolf/dhole genomes. It features scaffold, annotated gene and protein information.

Detailed tutorials are provided below:
- Browse data
  - Scaffold
  - Gene
- Search data
  - Simple Search
  - Blast
- Genome View
- Download

3. Dog SNP Database

Dog Genome SNP Database (DogSD) is a data container for the variation information of dog/wolf genomes. In the current version, DogSD contains 42 million non-redundant SNPs which curated from 127 individual samples and dbSNP (ver 146).

Detailed tutorials are provided below:
- Browse SNP data
  - SNP tables
  - SNPs in Gene
  - Browse on Chromosome
- Search data
  - Single individual search
  - Compared individuals search
- Download

5. Dog Expression Database

Dog Expression Database (DogED) is a repository of gene expression profiles derived entirely from RNA-Seq data analysis of tissues from Canis. Currently, 7 RNA-Seq Projects and 83 Experiments are collected from NCBI.

Detailed tutorials are provide below:
- Browse Data
  - RNASeq Project
  - Gene Differential Expression in Project
  - Annotated Gene
- Search Data
  - Gene Expression Search

6. Dog Human Disease Connection

Dog Human Disease Connection (DogHDC) is dedicated to provide homolog information for dog disease associated genes in other species, especially in human. All homolog information is curated from NCBI Homolog Gene as well as OMIM.

Detailed tutorials are provided below:
- Browse Data
  - Dog & Human homolog information
- Search Data
  - Dog & Human homolog information

7. Dog Online

Dog online provides data visualization and analysis tools for iDog. Users can analyze their own data and visualize the results online.

Detailed tutorials are provided below:
- Data Visualization
  - Dog SNP Genome Browser
- Data Analysis
  - BWA

FAQ

How to find interesting gene? FAQ

Dog Genome Database (DogGD)

1. How to search annotated information in DogGD? FAQ

2. How perform blast with my own data against DogGD? FAQ

Dog SNP Database (DogSD)

1. How to retrieve RefSNP information? FAQ

2. How to retrieve Individual SNP information? FAQ

3. How to find the SNP annotated information? FAQ

4. How to search for SNP in single individual? FAQ

5. How to search for SNP in multiple individuals? FAQ

6. How to download data of DogSD? FAQ

Dog Phenotype Database (DogPD)

1. How to search breed information? FAQ

2. How to search dog disease information ? FAQ

Dog Expression Database (DogED)

1. How to find gene profiles by FPKM value? FAQ

Dog Human Disease Connection (DogHDC)

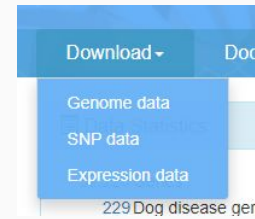1. How to find homolog information of dog and human? FAQ

Downloadable Data

- The downloadable files hold a lot of memory, which results in a slow download time
- Specific Programs are needed to run downloaded files (.fa and .bam extensions types)
- Converting these files is problematic

Demo

- Sparky is a 13 yr. Old Yorkshire Terrier
- Lately has developed cloudiness in his eyes
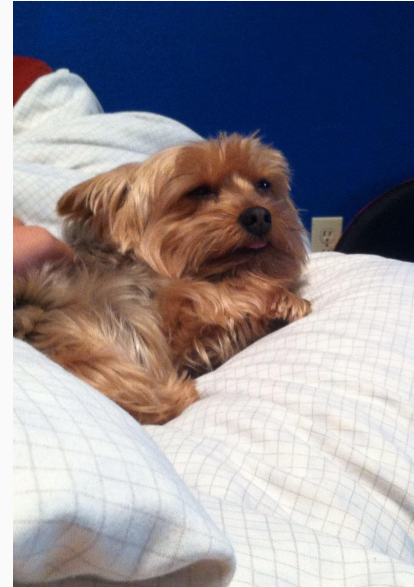
Judgement

- This database is not useful or easy to navigate if a person is not familiar with the field of study
- Data is mainly about dogs, but presented in a professional fashion
  - Curated from many credible databases
  - Published in credible journal
  - Mainly useful for scientists, geneticists, dog owners