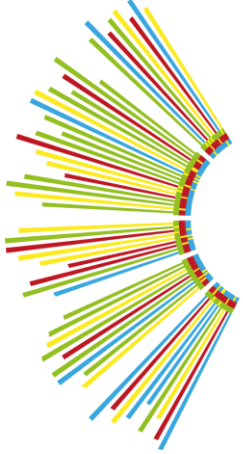


FAANG

Functional Annotation of Animal Genomes

# Data Management

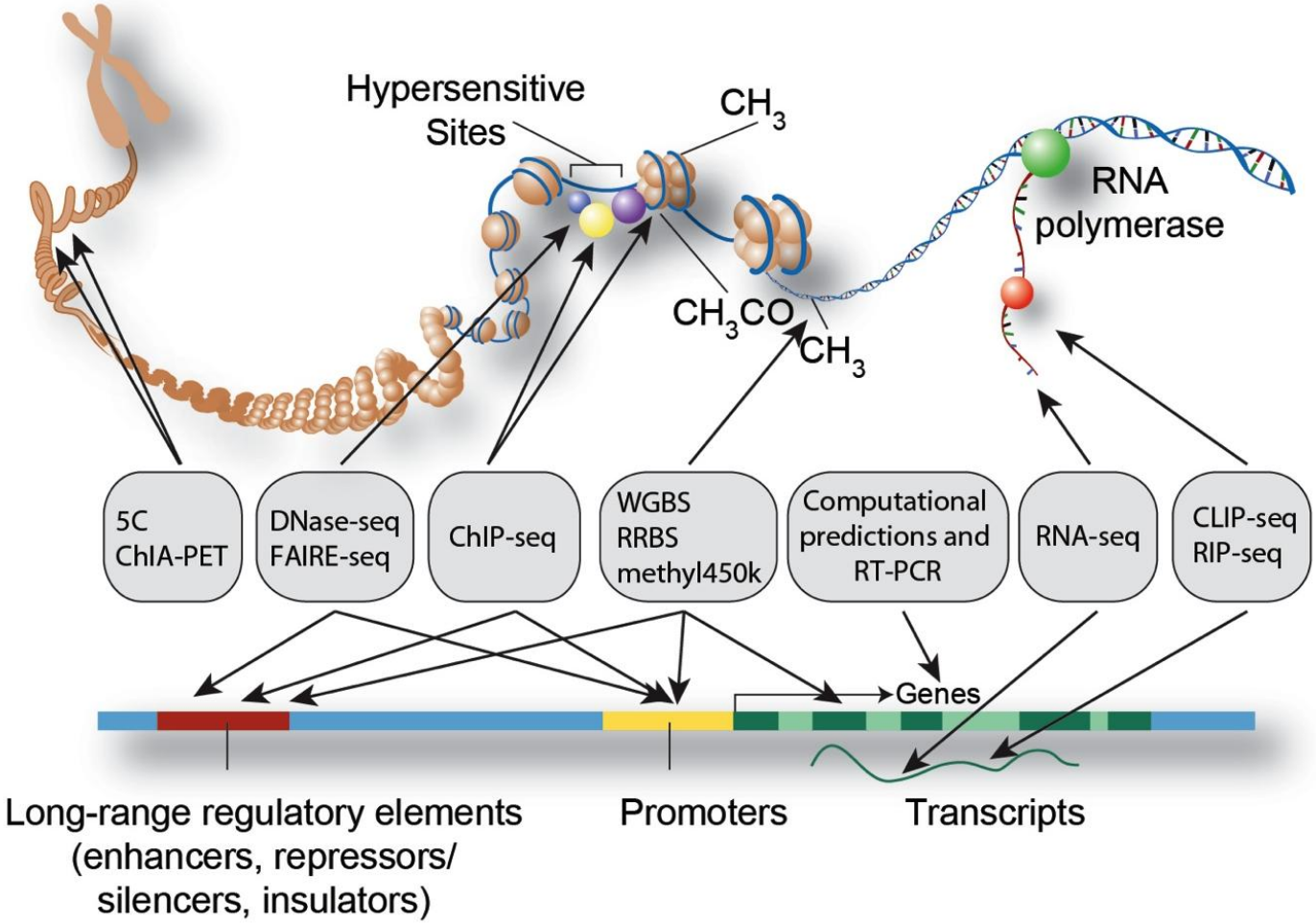
## Opportunities, Challenges and Connections to the Pilot Projects

Laura Clarke

6th February 2015

EMBL-EBI

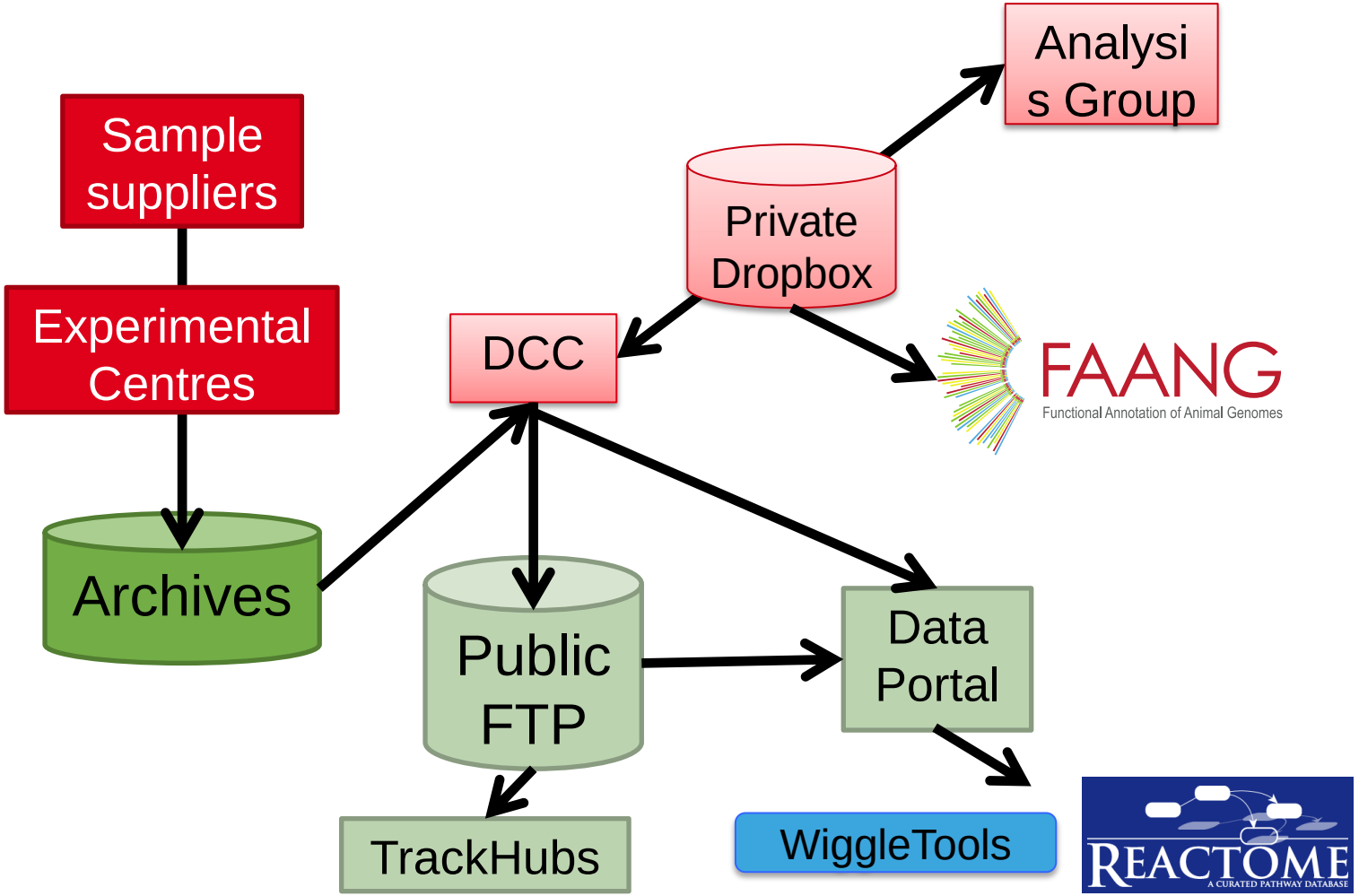
# Opportunities



# Challenges

- Data Flow
- Standards
- Communication
- Accessibility

# Data Flow


Sample suppliers
Experimental Centres
Archives
DCC
Private Dropbox
Analysis Group
FAANG
Functional Annotation of Animal Genomes
Public FTP
Data Portal
TrackHubs
WiggleTools
REACTOME
A CURATED PATHWAY DATABASE
e!Ensembl
UCSC

# FAANG Standards

- Animals, Samples  and Protocols
  - Elisabetta Giuffra, Huaijun Zhou
- Metadata and Data Sharing
  - Laura Clarke, Carl Schmidt
- Bioinformatics and Data Analysis
  - Jim Reecy, Mick Watson
- Communication
  - Chris Tuggle, Jeff Silverstein

# Animals, Samples and Protocols

- Who are collecting what samples
- Which breeds/lines?
- What assays are core
- Experimental replication?
- What other assays are done
- What are the assay standards
- Cross talk with the Bioinformatics group about metrics for quality
- Cross talk with Metadata about Ontologies

# Core Assays

- Transcriptomics
  - **Stranded RNA-seq** (exhaustive catalogues of gene expression; starting point for improving genome annotation)
- Chromatin Accessibility
  - **DNase I-seq (**DNAse hypersensitivity) or **ATAC-seq** (Transposase-Accessible Chromatin with high-throughput sequencing)
- Chromatin Modification
  - **H3K4me3** (active promoters/tss)
  - **H3K27me3** (silencing)
  - **H3K27Ac** (active regulatory elements)
  - **H3K4Me1**(distal regulatory elements and enchancers).

# Metadata Standards

- What metadata is essential
- Ontologies
- Validation of meta data
- Facilitate data sharing
- Tools for distribution
- Coordinate data releases
- Ensuring processes are transparent and reproducible
- Data Archiving support
- Data Formats
- Cross talk with Samples about Ontologies

# Bioinformatics Standards

- Uniform analysis pipelines
- ? Benchmarking of different groups pipelines
- Standard Reference Datasets
  - Genome
  - Gene set
- Minimum aligned coverage
- Standard normalization methods
- File Formats
- Reference IDs for collections of data

# Communication

- Lab Exchanges
- Wiki/Intranet for document exchange
  - http://www.faang.org/wiki
- Promotion of FAANG with funding agencies
- Social media to engage with the community
  - @faangomics
- Organise in person meetings

# Accessibility

- Frequent Data Release
- Submit to archives often and early
- FTP sites
- TrackHubs
- Portal to aid discoverability
- Interactive Tools
  - WiggleTools

# Accessibility - Ensembl Regulation

The goal of Ensembl Regulation team is to annotate the genome with features that may play a role in the transcriptional regulation of genes.

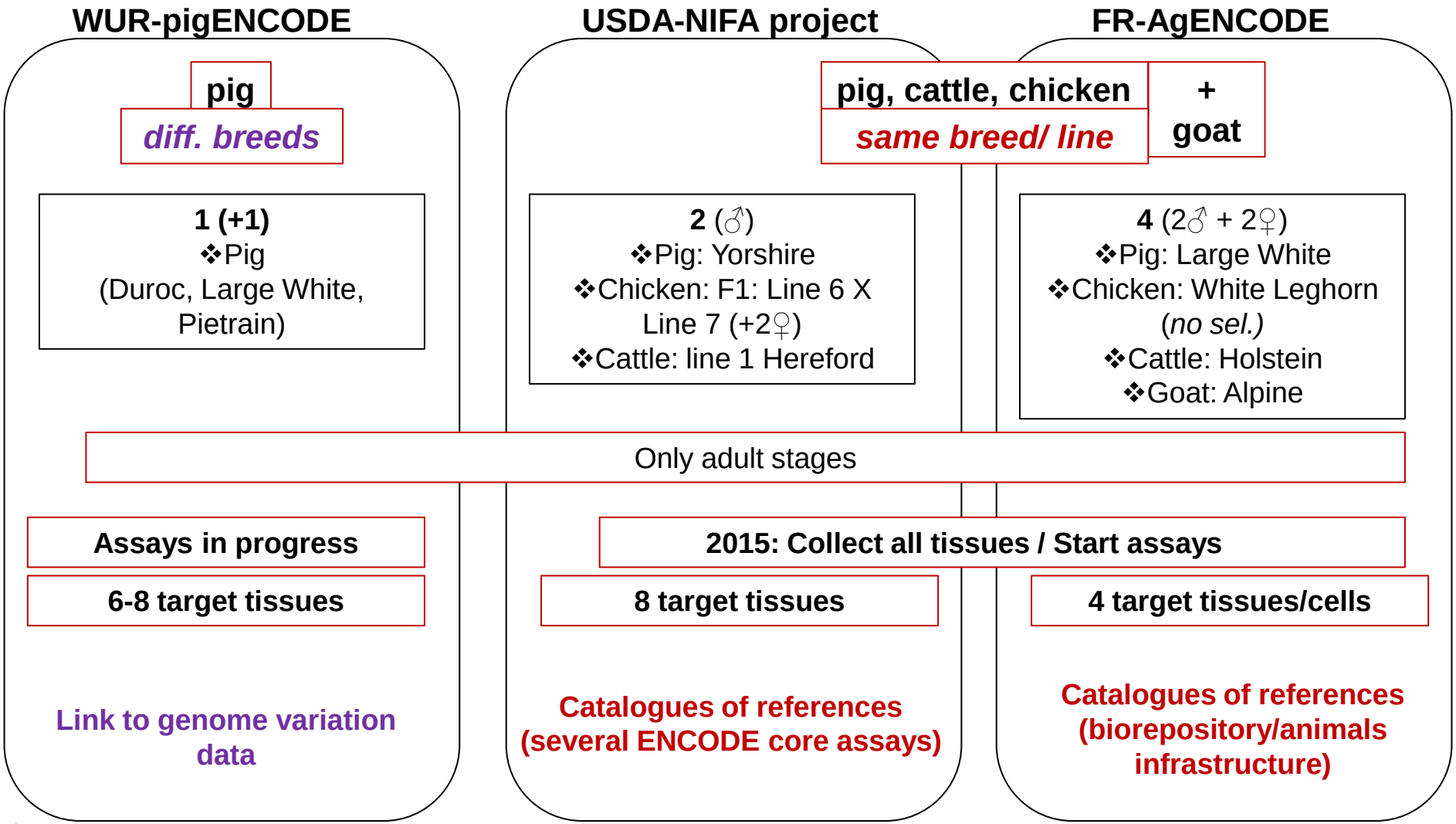
- Predicted open/closed chromatin
- Transcription factor binding sites
- DNase I sensitivity
- Epigenetic marks
- RNA Pol binding

# Initial Projects

**Common aims: improve the functional annotation of the genomes of major domesticated species. *working in the FAANG framework***

- **WUR-pigENCODE**
  - Wageningen University, **Martien. Groenen** Univ. Illinois (US) and INRA (France): E. Giuffra., *Funding:* ERC-grant Started: **01/01/2014**
- **USDA – NIFA project**
  - UC-DAVIS: **Huaijun Zhou**, USDA, ARS ADOL; Iowa State University: Michigan State University: USDA ARS, Miles. *Funding:* USDA, Chicken, Swine, Bovine Species Genome Coordination Funds, National Pork Board, Aviagen Started: **01/01/2015 (36 months)**
- **Fr-AgEncode**
  - INRA: **Elisabetta Giuffra***:* INRA *Funding:* metaprogramme on Genomic Selection (SelGen) Started: **01/01/2015 (30 months)**
- **FAANG UK BBSRC slola (under review)**
  - Roslin, **Alan Archibald** TGAC and EMBL-EBI. *Funding* BBSRC slola (if granted) **hopes to start 10/2015**

# Biological Targets and Resources

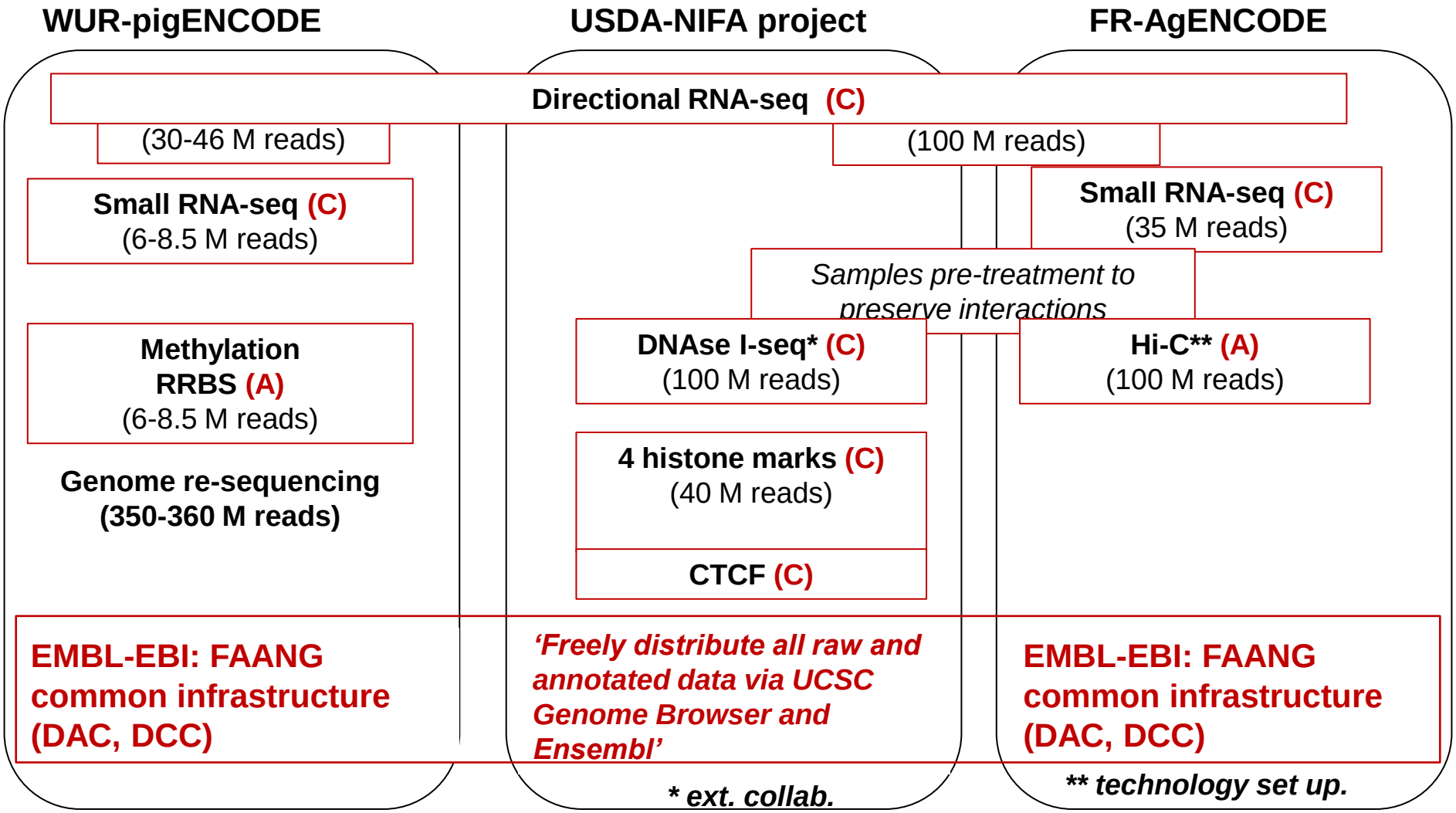



FAANG Functional Annotation of Animal Genomes

http://www.faang.org

EMBL-EBI

# Choice of FAANG assays: C (core), A (additional)




FAANG Functional Annotation of Animal Genomes

http://www.faang.org

EMBL-EBI

# Please join us at http://www.faang.org

# Thanks

# Questions?