

# *Functional Annotation of ANimal Genomes (FAANG)--*

## *Global FAANG Consortium*

Organic growth of FAANG during 2014-2019

Current FAANG contributors= 432



2014 membership- GB authors



March 2019 membership



# Outline – Introduction to FAANG Consortium

1. Why FAANG?
2. ENCODE results as a FAANG goal
3. Brief history and goals
4. FAANG Committees and activities
5. Summary of current FAANG projects
6. FAANG data available → getting involved

# Why is FAANG important?

- Understanding the genotype to phenotype link:
  - Providing information to accelerate genetic improvement
  - Improving fundamental understanding of biology
- We need to ***annotate the function*** of the genome of our animals

# What is functional annotation?

- We have the DNA sequence for the genome of our major domesticated animals...

AGGGCCGTTGGGCTCCTCTGATTGGACGTGGCTGCGCCTGGCTGGCCAATGGGGAGATGGCGGCAGGTTCCCGGAAGTGGTGGCGCGGTGAGCTGCAGCTTGCATAACC  
GCGGAGGGCTGGCGCCCAGTCGGCTGTTCGCCAACGCTATGAGTTCCTTCGAGGGCCAGATGGCGGAGTACCCAATCTCCATAGACCGTTTCGACCGGGAGAATC  
TGAGGGCTCGCGCTTATTTCTGTCCACTGCCACAAGGTTGAGTGAGCGCGGCGCGCCGACCCCTCCCGGGGACCAGGGCTGCGGCGGGTCTGGCCCCGCGGGGAG  
GTAGCCAGGGGCTGGAGAGGAGGAAGTTGGGGTGGGGTCTGCAGGAGGAGTGCAGGCTTTGGGAGGTCTGGGGAGAGACTAGCGACTGCATCCTGTGCGAGGC  
GTGCTCTTCAGCCTGTTTGAGGCTTTATCGCCTC  
TTTCGTTTTTTTTGGAAGCTCTCCAGGGCCCAAT  
CTGCCATACCTCCAACCATTTATTAGCAAGCACG  
CAGTATGAAAGTTCTTTTTGTCCTAGAGGTGGC  
CTGGAATGTCGCCCCCTAACGATATGGTCCCCA  
CCATTTCCATTCTTTTTTTTTTTTTTTTTCTTTTTG  
CACCATAGCCACAGCAACTCAGGATCCAAGCCG  
ACCTCATGGTTCCTAGTCGGATTCATTAACCACT  
AAACAGGTTTAATATTAGTATGCTAATTAGCAC  
AGCTTAAGTAGGAGACCCCGACCGCTCTTTTGA  
CTTGGGCAGGCAGTGGGTGTGAAGACTAGCTGC  
ACTCTTTAATGAGGAAATGCTGGACCTCCTCTCT  
TTTTATTGAGAAGTCATTCCTGGAAGATGGTGT  
TTCCATTGTGGCGCAGTGGAATGAACCCGAC  
TGTAGGTCACAGACATTGTTGTGCTGTGGCATA  
AAAAAGACAAAAAAGAAAAAAGAAAAAAGGAA  
AGGAGTTCGATTTATATTTTTATTGAATTTATT  
ATAACTTTCTGTCTCTGTTCCCATGTGGGCC  
GTGGTTTTAACTTTTTGGGGAAGCTGACGTCCTTGGATGTGCTGTGGGCTGGCTTGTGTCTCTCCCTACCCCATTCATTTTTGAATCCCTAATCTCCGATGTGGCTGTA  
TTTGGAGAAAGGGCTTTTAGGAGGTGATGGAGGTTAAATGAGGTCATAAGGGTGGAGTCTAATTTTATAGGAGTGGTGGCTTCTCTACACACACACACACACACAC  
ACACACACACACACTGAGGAATGCCGTGTAGGCACGCGCATGAAAGTGGCTGCTGGTAAGGAGAGGCCTCCCGAGAAACCAGCTTTGGTAGCTTGATCTTGGACTT  
GAGCTTCCAGAACCGTTAGAAATACATTTCTGGGTGTTTGCAGTGTGGCACAGTGGGTTAAAAACCTGACTGCAAGTGGCTCAGATCTCTGTAGCCGTGTGGGTTCCGTCC  
CTGGCCAGTGCAGTGGGTTAAGGATCCAGCCTTGCCACAGCTGTGGTGTAGGTTGCAGCTGCAGCTTGGAGTCAATCCCTGEGCTGGGAAAGTTCCATATGCTGCAGGGA  
CAGCGGTAGGAAAAAACAATAAACAATAAATAAATAAATAAAGGAGTTCCTATTGTGGTACAACAGGGATGAATCCAATAGGAACCATGAGGTTGCAGGTTGATCCC  
TGGCCTCGCTCAGTGGGTTAAGGATCCGACATTGCCATGAGCTGTGGTGTGGGTCGAAGATGAGGCTCAGATCTGGTGTGCTATGGCTGTGGTGTAGGCTGGCAGCTG

This is not English!

What is important information and what isn't??

# Annotation can help....

- Find what part codes for proteins
- Understand effects of genetic variation on these proteins

1	AGATGGCGGCAGGTTCCCGGAAGTGGTGGCGCGGTTCAGCTGCAGCTTGCATAACCGCGGA	60
61	GGGCTGGCGCCCAGTTCGGCTGTGTTCGCCAAGCTATGAGTTCCTTCGAGGGCCAGATGG	120
121	CGGAGTACCCAACCTATCTCCATAGACCGTTTCGACCGGGAGAATCTGAGGGCTCGCGCTT	180
181	ATTTCTGTCCCACCTGCCACAAGGATCACATGAAAGGATTAAGAGCCTCTACCTTGAAAA	240
241	GAAGTTGGAGTGCAGCTTGAAGTCTCCTTATACTGTTACCTGTTACTAGAGAATTGT	300
301	TATTAACCAACCCGAGGTACAGATTTGGGAGAAAAGCAATTGTCTCAATTGAAGTTGAAA	360
361	CTCCTACCCAGATATCTTTAATTGATGAAGCATCAGGCGAGAAGGAAGAAATTGTTGTGA	420
421	CTCTTTACCAGCTGGTCATTGCCAGGATCAGTTATGTTCTTATTTTCAGGGCAACAATG	480
481	GAACTGTCTTGTATACAGGAGACTTCCGATTGGCAAAAAGGAGAAGCTGCCAGAATGGAGC	540

.....

26	CGGAGTACCCAACCTATCTCCATAGACCGTTTCGACCGGGAGAATCTGAGGGCTCGCGCTT	85
86	ATTTCTGTCCCACCTGCCACAAGGATCACATGAAAGGATTAAGAGCCTCTACCTTGAAAA	145
146	GAAGTTGGAGTGCAGCTTGAAGTCTCCTTATACTGTTACCTGTTACTAGAGAATTGT	205
206	TATTAACCAACCCGAGGTACAGATTTGGGAGAAAAGCAATTGTCTCAATTGAAGTTGAAA	265
266	CTCCTACCCAGATATCTTTAATTGATGAAGCATCAGGCGAGAAGGAAGAAATTGTTGTGA	325
326	CTCTTTACCAGCTGGTCATTGCCAGGATCAGTTATGTTCTTATTTTCAGGGCAACAATG	385
386	GAACTGTCTTGTATACAGGAGACTTCCGATTGGCAAAAAGGAGAAGCTGCCAGAATGGAGC	445

.....

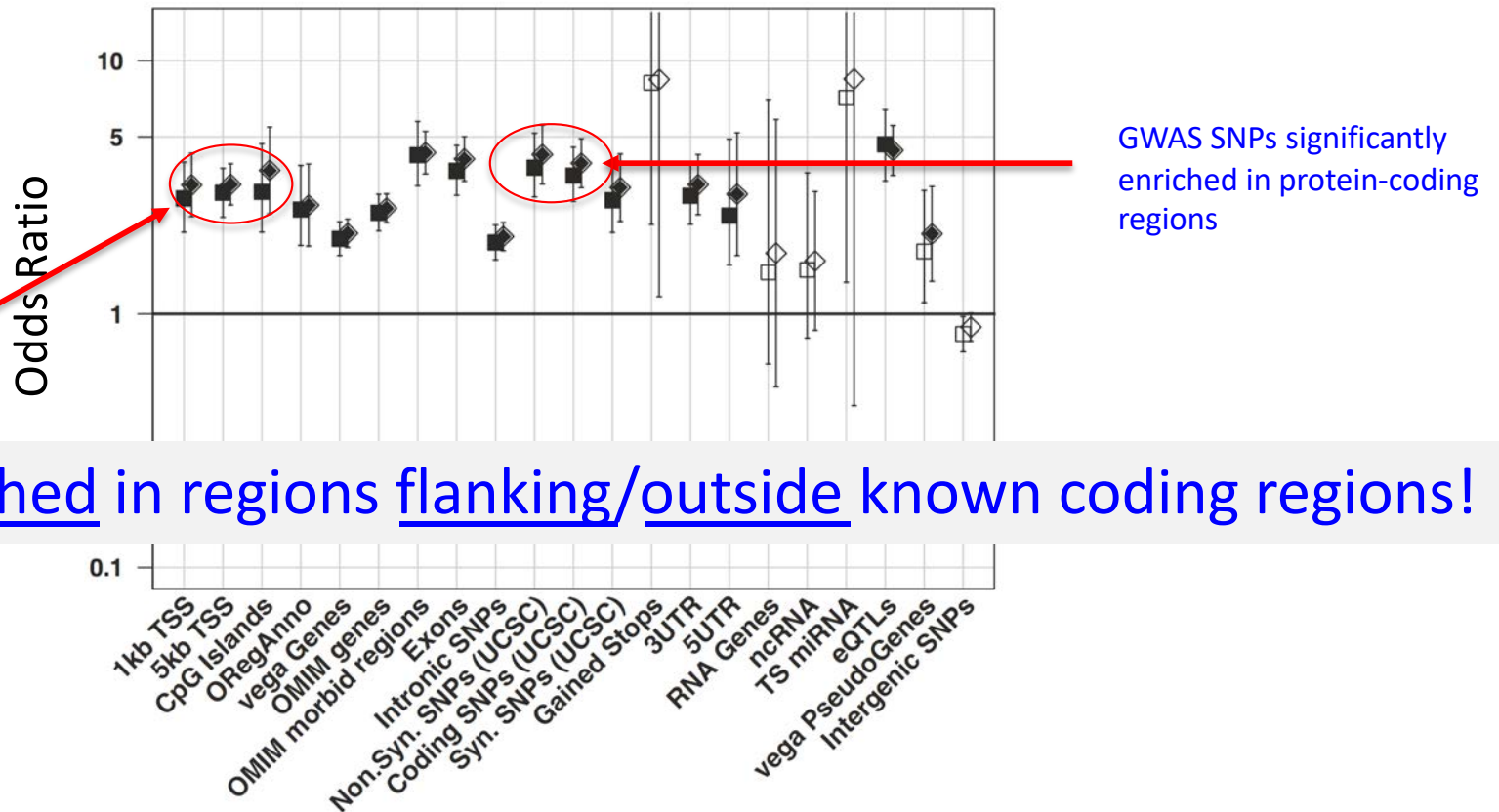
9	A--E--Y--P--T--I--S--I--D--R--F--D--R--E--N--L--R--A--R--A--	28
29	Y--F--L--S--H--C--H--K--D--H--M--K--G--L--R--A--S--T--L--K--	48
49	R--R--L--E--C--S--L--K--V--S--L--Y--C--S--P--V--T--R--E--L--	68
69	L--E--T--N--P--R--Y--R--F--W--E--K--R--I--V--S--I--E--V--E--	88
89	T--P--T--Q--I--S--L--I--D--E--A--S--G--E--K--E--E--I--V--V--	108
109	T--L--L--P--A--G--H--C--P--G--S--V--M--F--L--F--Q--G--N--N--	128
129	G--T--V--L--Y--T--G--D--F--R--L--A--K--G--E--A--A--R--M--E--	148



# Why is functional annotation important?

Is trait-associated genetic variation located (enriched) in genes?

Yes!



- Permutations
- ◇ Random Sampling
- | 95% CI
- Significant Odds Ratios



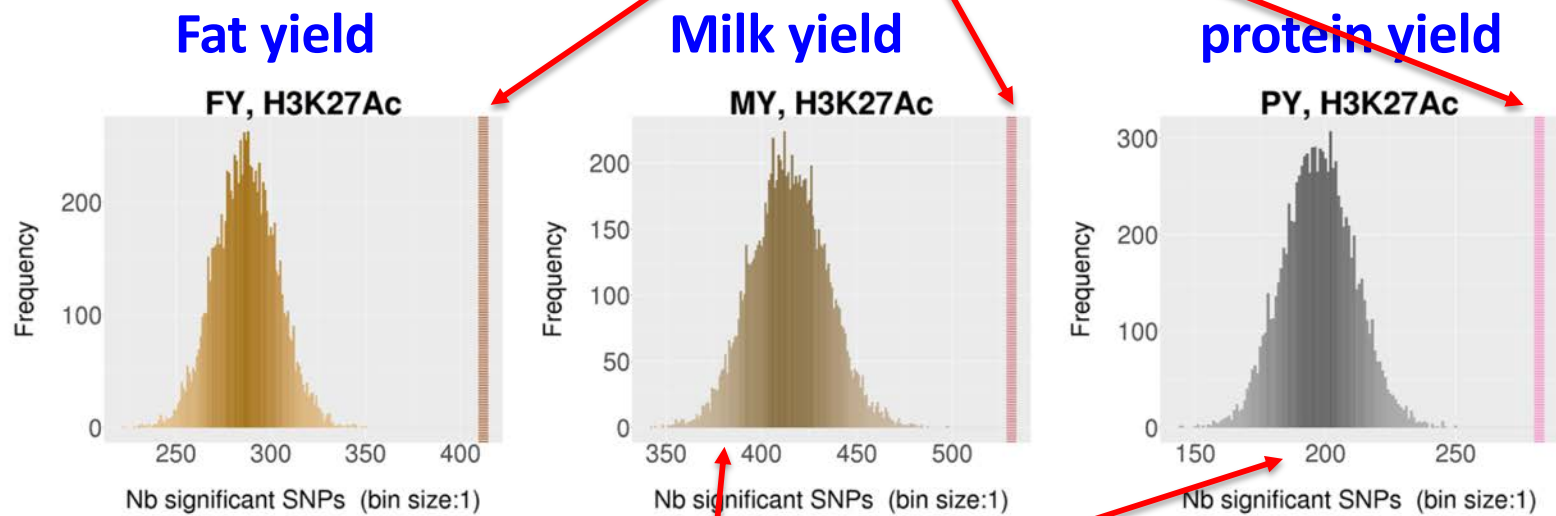
## The genomic signature of trait-associated variants

Alida S D Kindt, Pau Navarro, Colin A M Semple and Chris S Haley\*

BMC Genomics 2013 14:108

Shown in livestock as well: Milk trait-associated SNPs enriched in predicted regulatory regions

Many sig SNPs in bovine enhancer regions predicted from mapping ENCODE to bovine genome



Much fewer sig SNPs in random sampling

# Why is functional annotation important?

- SNPs associated with disease/traits are often outside genes  
i.e., Maurano et al. 2012 (human), Wang et al. 2017 (bovine)
- **This says the function of genomes are not just in proteins...**
- So we must understand function of “non-genic” regions of genome!
  
- But the genome is mostly “empty”- only 1.5% is protein-coding- which is the easiest parts to find and annotate...



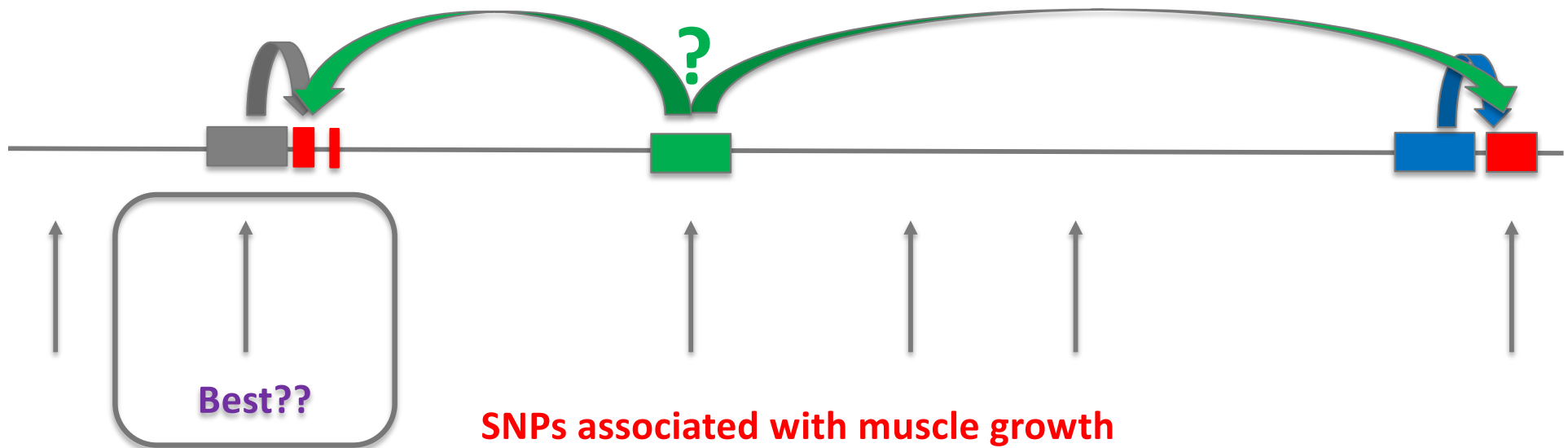
# How do we figure out what parts are important then?

## Protein-coding genes

Regulatory region- "ON" in skeletal muscle

Regulatory region- "ON" in mammary gland

Regulation- ???

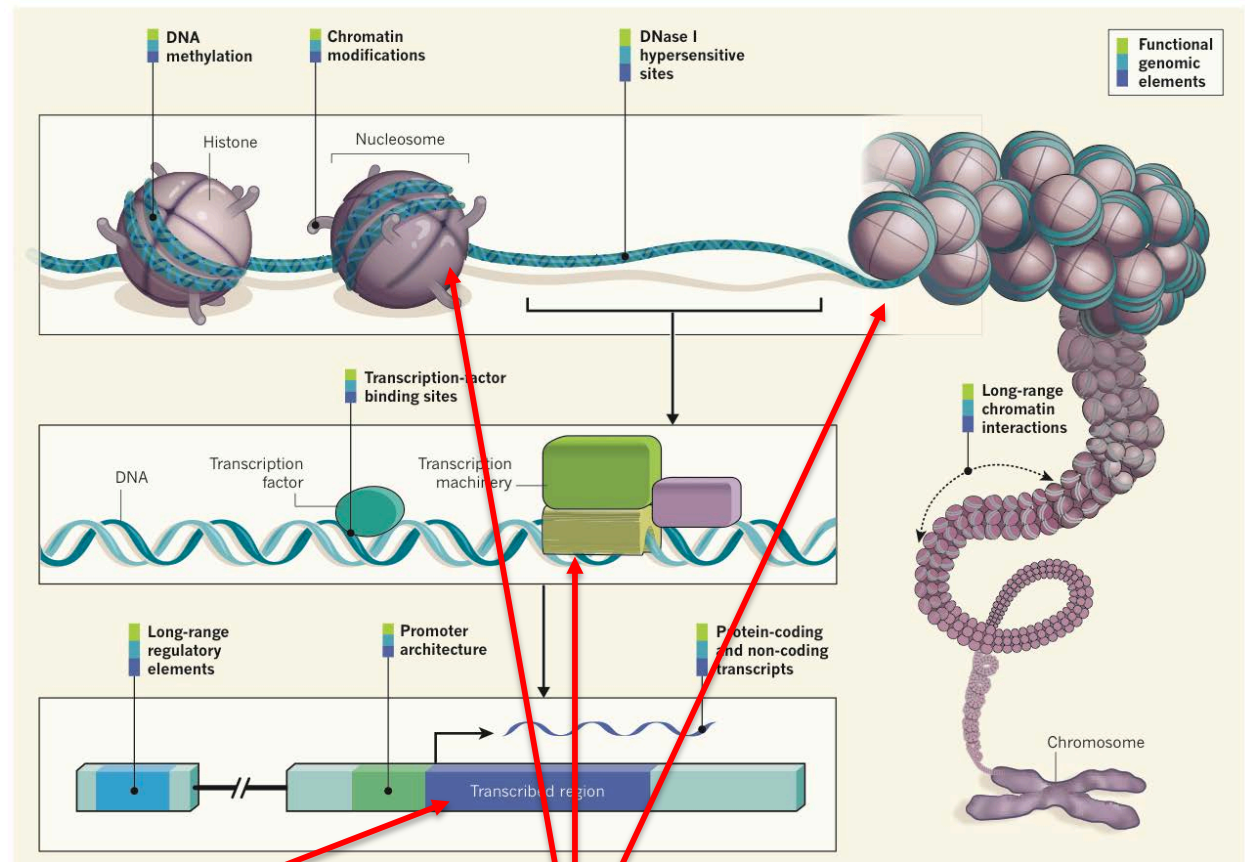


# Why is functional annotation important?

- If we want to be able to better predict trait value from understanding the importance of associated SNPs, we need to find all the functional parts of animal genomes
- *“Predicting phenotype from genotype”...*
- How to get there?

# Genome functional annotation described by ENCODE

- To understand function:
- a) what part is expressed?
  - b) what part controls this expression (specific tissue, response to infection, etc):

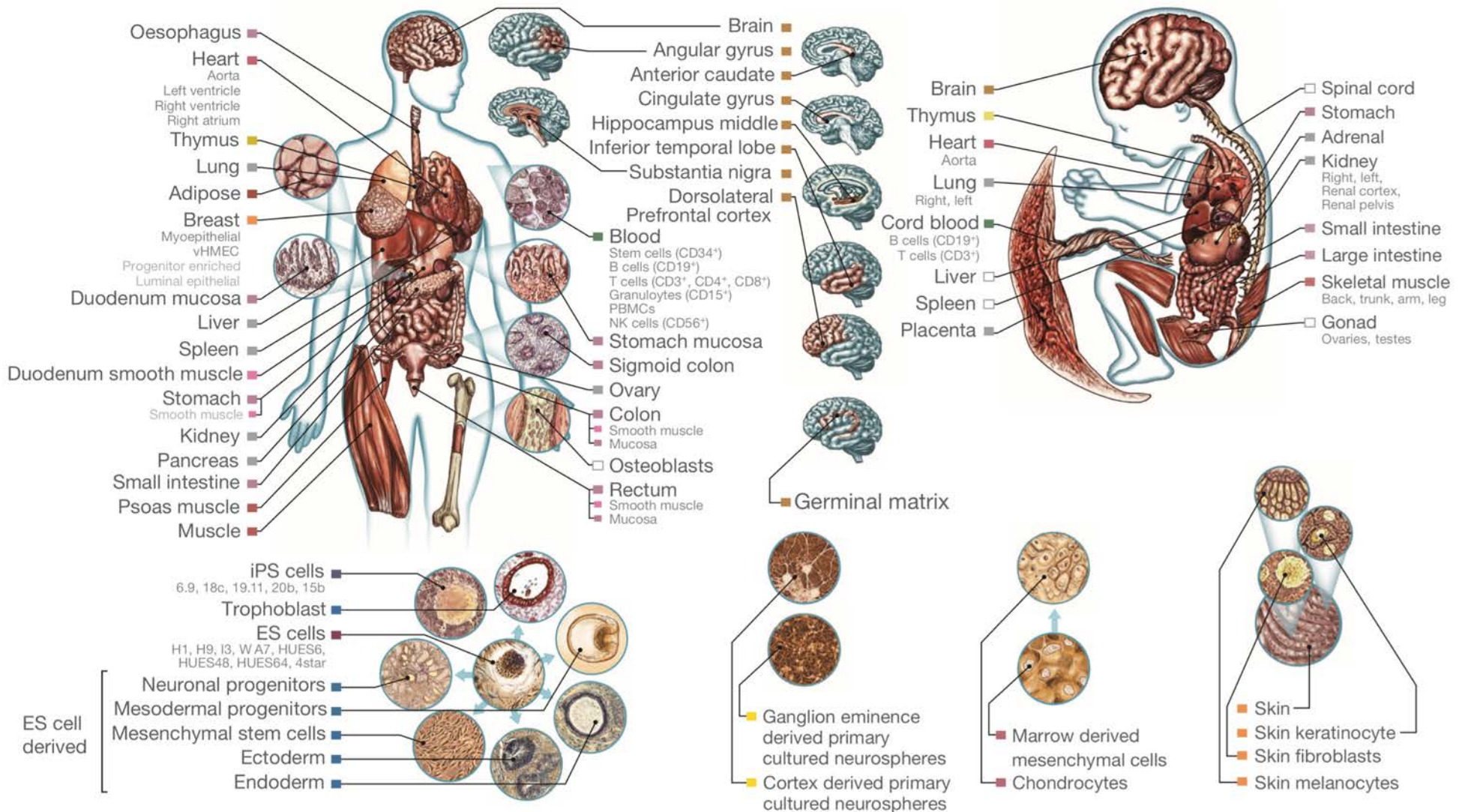


RNA levels and location

Biochemical assays of Chromatin structure

# Roadmap Epigenome: Extending ENCODE to tissues → many biological states

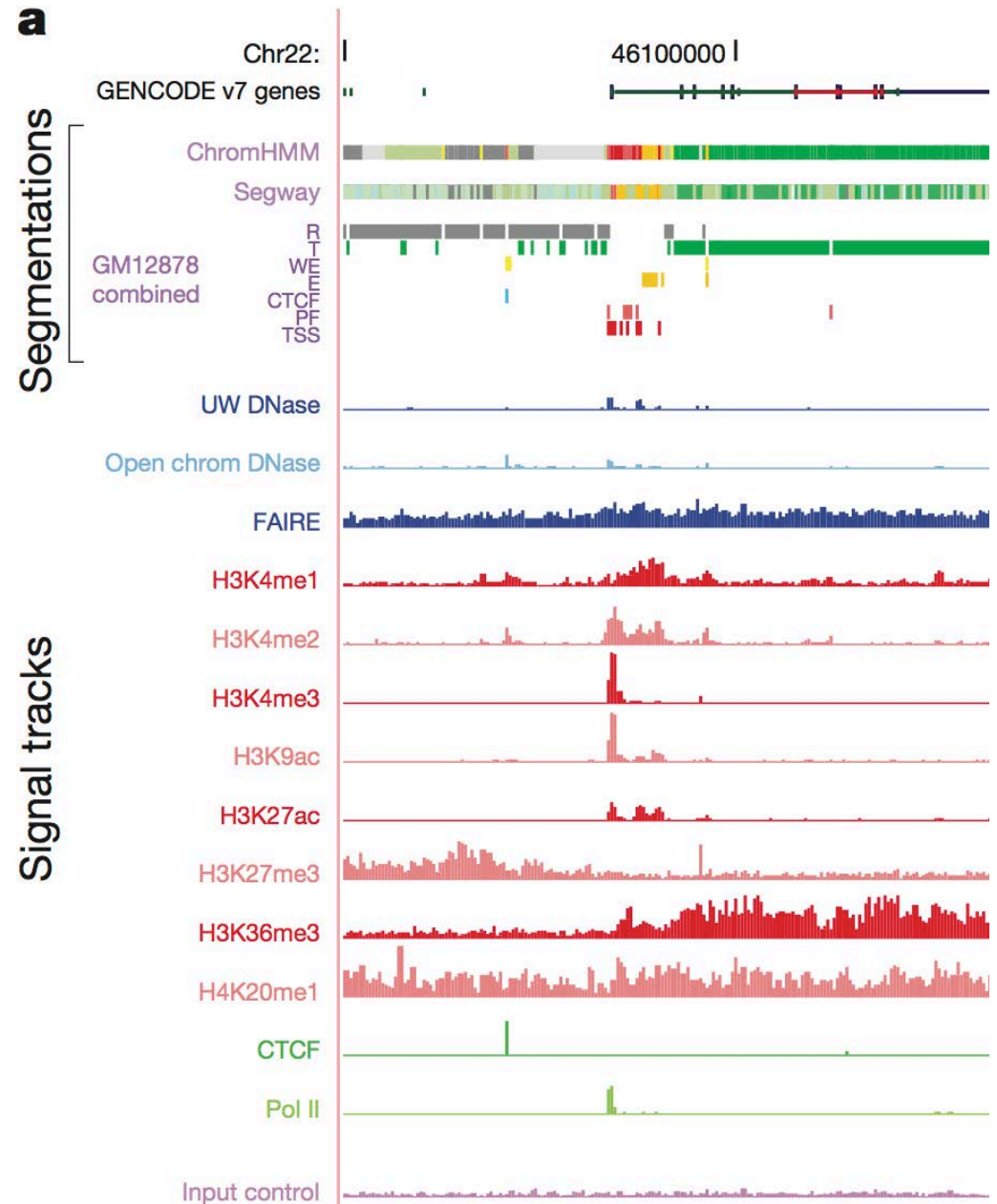
More biological states increases power to identify specific functional regions



# Functional Annotation results

Collecting all these data on the same samples allowed *predictive models for genome function* to be developed

- Predicts Chromatin State: OPEN or CLOSED
- Created a functional map of genome segments



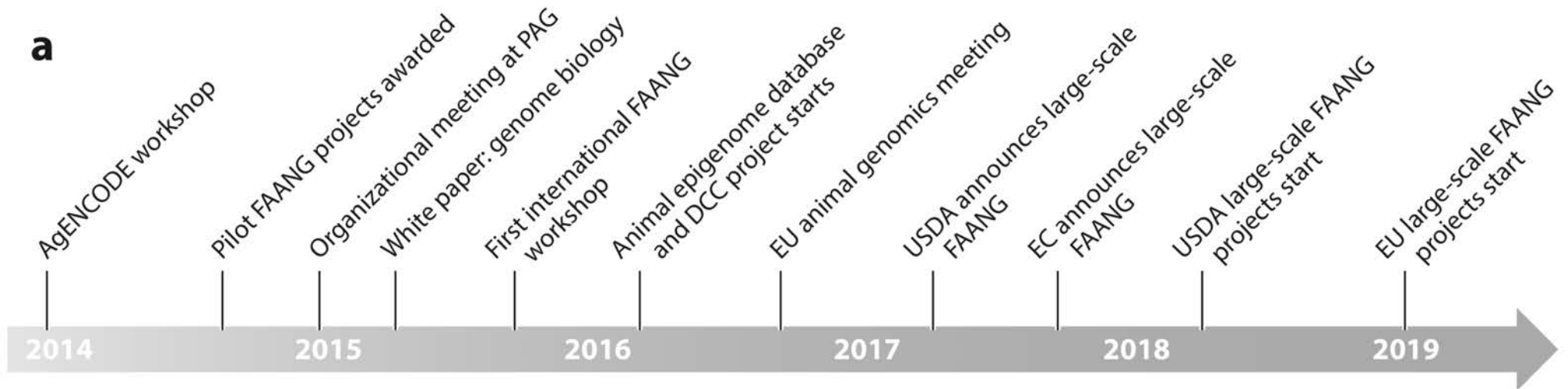
# Success in ENCODE and Roadmap required:

- High quality reference genome sequence
- Common infrastructure providing
  - Biological resources
  - Bioinformatics tools
  - Databases
- Effective coordination and communication

***FAANG needs all these to succeed!***

# Brief FAANG history


- Workshop in January 2014
- Organizational meeting in Jan 2015 in San Diego
- White paper published in 2015-- “Highly accessed”
  - > 9,400 accesses as of March 2019, [82 citations](#)
- [2015 Workshop to kick off FAANG](#)



Giuffra, Tuggle, et al., *Annual Reviews in Animal Biosciences*  
February 2019

# ***Gathering On FAANG Workshop October 7-8, 2015 National Academy of Sciences Building Washington, DC***

- 100 attendees
- Plenary talks and Breakout sessions
- **6 funding agencies presented talks**
- Meeting Report Published OA
- Set in motion organizing **calls for funding** for FAANG projects in US and Europe
- **RFP in 2017 at USDA, 2018 in EU**

**ANIMAL GENETICS** Immunogenetics, Molecular Genetics and Functional Genomics 

doi: 10.1111/age.12466

**GO-FAANG meeting: a Gathering On Functional Annotation of Animal Genomes**

Christopher K. Tuggle\*, Elisabetta Giuffra<sup>†</sup>, Stephen N. White<sup>†§¶</sup>, Laura Clarke<sup>\*\*</sup>, Huaijun Zhou<sup>††</sup>, Pablo J. Ross<sup>††</sup>, Hervé Acloque<sup>‡‡</sup>, James M. Reecy\*, Alan Archibald<sup>§§</sup>, Rebecca R. Bellone<sup>¶¶</sup>, Michèle Boichard<sup>†</sup>, Amanda Chamberlain<sup>\*\*\*</sup>, Hans Cheng<sup>†††</sup>, Richard P.M.A. Crooijmans<sup>‡‡‡</sup>, Mary E. Delany<sup>††</sup>, Carrie J. Finno<sup>§§§</sup>, Martien A. M. Groenen<sup>‡‡‡</sup>, Ben Hayes<sup>¶¶¶</sup>, Joan K. Lunney<sup>\*\*\*\*</sup>, Jessica L. Petersen<sup>††††</sup>, Graham S. Plastow<sup>‡‡‡‡</sup>, Carl J. Schmidt<sup>§§§§</sup>, Jiuzhou Song<sup>¶¶¶¶</sup> and Mick Watson<sup>§§</sup>





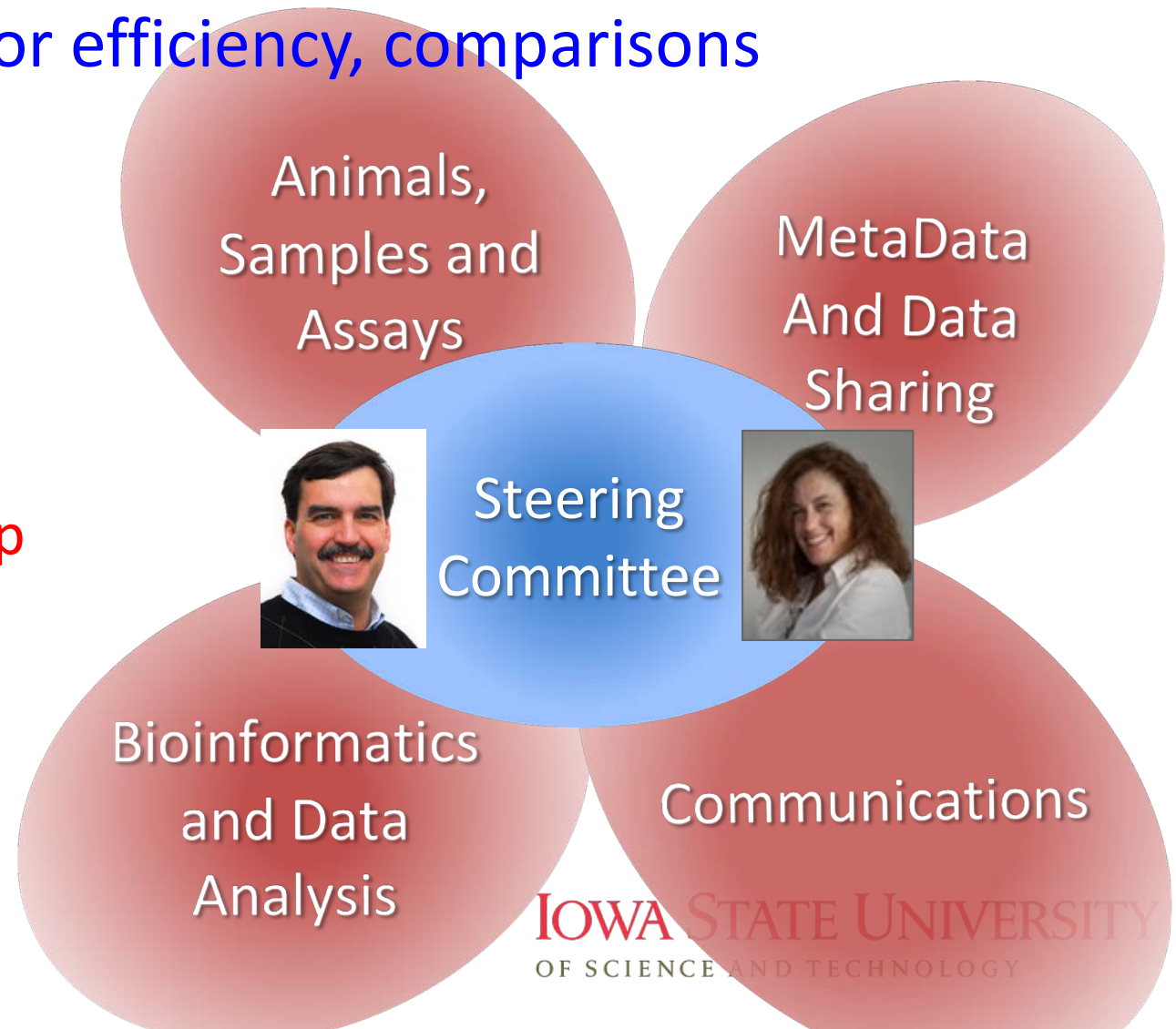
# FAANG is global- >400 members across six continents

- FAANG will globally coordinate functional annotation across species for efficiency, comparisons

FAANG Committees have conference calls on-line periodically

FAANG members can self-select membership in these committees

**Anyone can join FAANG!**



# *Main current activities*

- **Establish set of Core assays, begin to develop ENCODE-type functional data**
- **Develop tissue description, storage and sharing protocols**
- **Develop computational tools to analyze data**
- **Develop bioinformatics infrastructure**
- **Develop communication mechanisms**

Reviewed in *Annual Reviews in Animal Biosciences* February 2019

Giuffra, Tuggle, et al.

## Functional Annotation of Animal Genomes (FAANG): Current Achievements and Roadmap

### Annual Review of Animal Biosciences

Vol. 7:65-88 (Volume publication date February 2019)  
First published as a Review in Advance on November 14, 2018  
<https://doi.org/10.1146/annurev-animal-020518-114913>

Elisabetta Giuffra,<sup>1</sup> Christopher K. Tuggle,<sup>2</sup> and the FAANG Consortium\*

<sup>1</sup>Génétique Animale et Biologie Intégrative (GABI), Institut National de la Recherche Agronomique (INRA), AgroParisTech, Université Paris Saclay, 78350 Jouy-en-Josas, France; email: [elisabetta.giuffra@inra.fr](mailto:elisabetta.giuffra@inra.fr)

<sup>2</sup>Department of Animal Science, Iowa State University, Ames, Iowa 50011, USA; email: [cktuggle@iastate.edu](mailto:cktuggle@iastate.edu)

# *Main current activities*

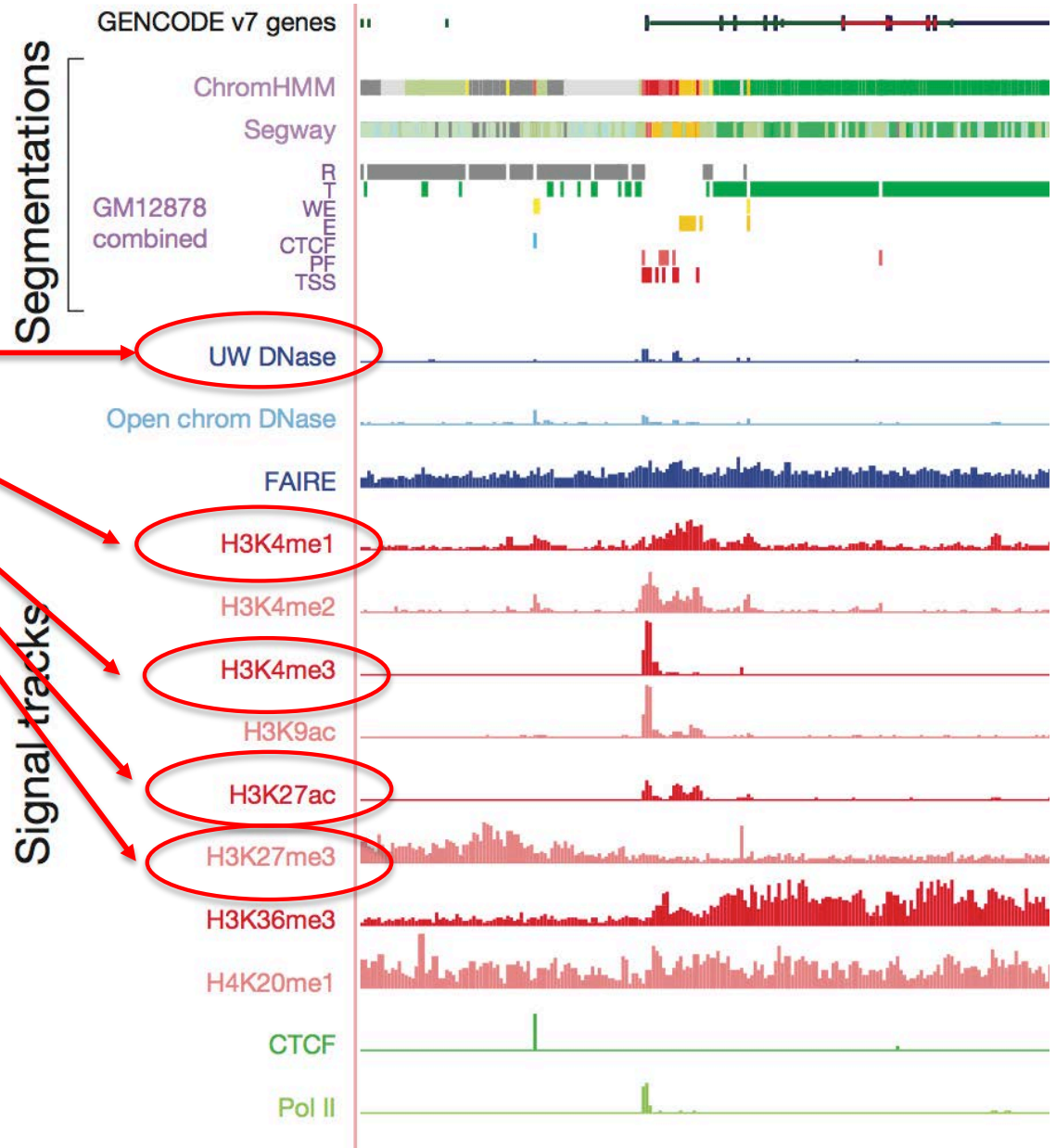
- **Phase I**
  - **Emphasis on establishing reference datasets for healthy adults**
  - **Few biological replicates, no treatments**
  - **Validate metadata toolboxes and analytical pipelines**
- **Phase II**
  - **Expand biological states- treatments, developmental stage, genetic variation**

# *Core assays needed for chromatin state modeling*

- **RNAseq (stranded, deep) analysis of many tissues from the same 2-4 individuals**
- **Chromatin accessibility/architecture assays on these tissues (DNase I or ATAC-seq)**
- **Histone modification marks**
  - **H3K4 monomethylation**
  - **H3K4 trimethylation**
  - **H3K27acetylation**
  - **H3K27 trimethylation**

# Minimum information for finding functional regions in livestock genomes

**FAANG core assays**



## *Core assays needed for modeling*

- RNAseq (stranded, deep) analysis of as many tissues from the same 2-4 individuals
- Chromatin accessibility/architecture assays on these tissues (DNase I or ATAC-seq)
- Histone modification marks
  - H3K4 monomethylation
  - H3K4 trimethylation
  - H3K27acetylation
  - H3K27 trimethylation

## *Additional assays to determine genome function*

- **DNA methylation**
- **Regulatory factor protein binding (CTCF)**
- **Chromatin interaction (Hi-C)**

# FAANG projects- an early list

Species	Leading Inst. & country	FAANG- FAANG-	
		<i>pilot</i>	<i>rel.</i>
pig, cattle, goat, chicken	INRA	France	<b>Sylvain Foissac talk today</b>
pig, cattle, chicken	UC-Davis	US	X x
horse	UC-Davis & Nebraska Univ.	US	<b>Jessica Peterson talk today</b>
cattle	Leibniz Inst. for Farm Anim. Biology	Germany	X x
Sheep	Int. Sheep Genome Consortium & others	Australia & US	<b>Brenda Murdoch talk today</b> X x
pig, chicken	Wageningen Univ.	The Netherlands	X X
cattle	Dairy Futures Coop. Res. Centre, AgriBio & others	Australia	X X
cattle, pig	Alberta & Guelph Univ. & others	Canada	x X
cattle and zebu	Adelaide Univ.	Australia	x X
sheep, buffalo	Roslin- Edinburgh Univ.	UK	x X
cattle	Washington State Univ.	US	x X
'Arctic Ark'	Natural Resources Institute	Finland	X

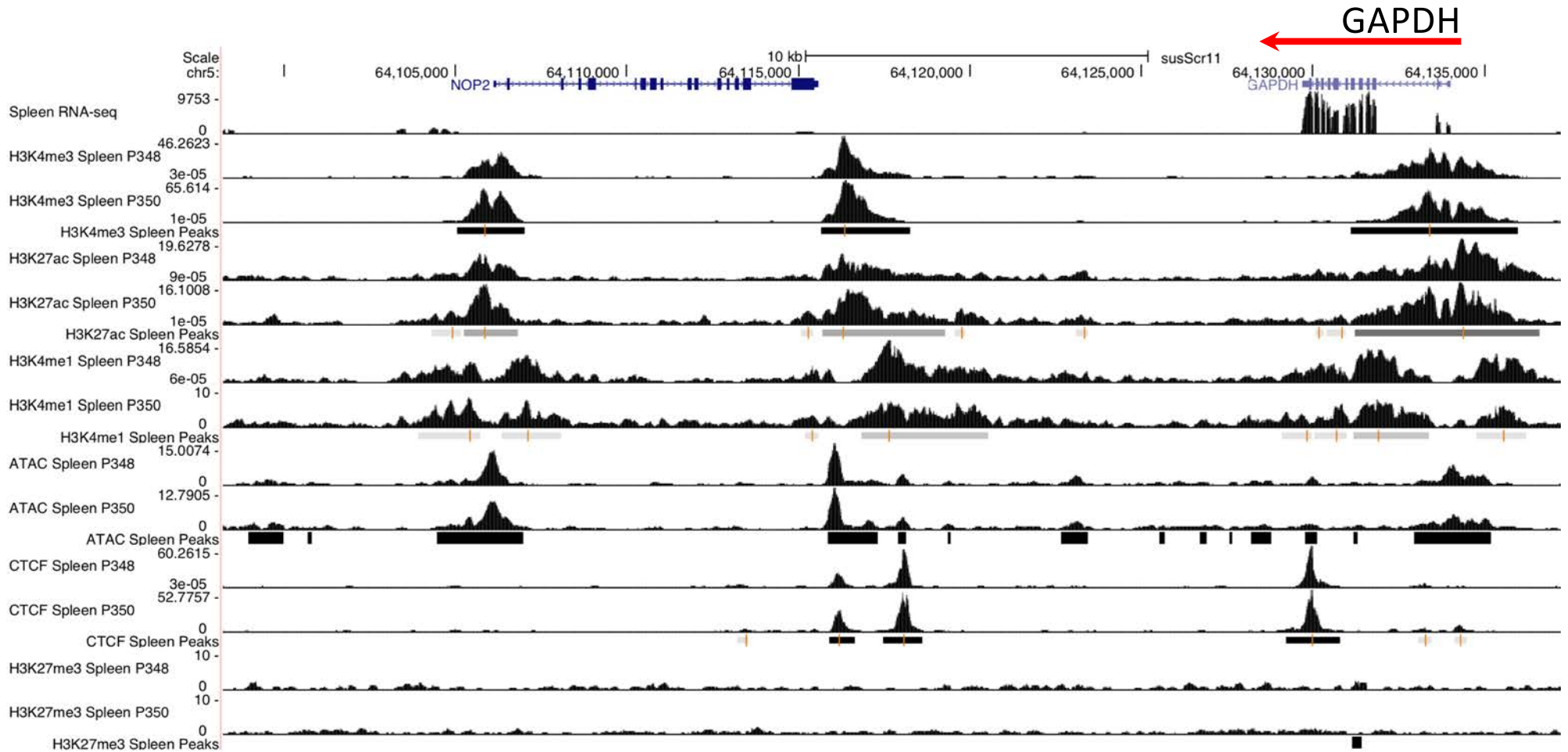
# Progress on UCD FAANG Pilot project



<b>RNA-seq</b>	Data Analysis		Data Analysis		Data Analysis
<b>DNase/ATAC-seq</b>	Data Analysis		Data Analysis		Data Analysis
<b>H3K4me3</b>	Data Analysis		Data Analysis	Chromatin Sheared	Data Analysis
<b>H3K27me3</b>	Data Analysis		Data Analysis	Chromatin Sheared	Data Analysis
<b>H3K4me1</b>	Data Analysis	Chromatin Sheared	Data Analysis	Chromatin Sheared	Data Analysis
<b>H3K27ac</b>	Data Analysis	Chromatin Sheared	Data Analysis	Chromatin Sheared	Data Analysis
<b>CTCF</b>	Data Analysis	Chromatin Sheared	Data Analysis	Chromatin Sheared	Data Analysis
<b>RRBS-seq</b>	Data Analysis		Data Analysis		Data Analysis
<b>WTTS-seq</b>	Data Analysis		Tissues Collected		Tissues Collected



# Example data for UCD pilot FAANG



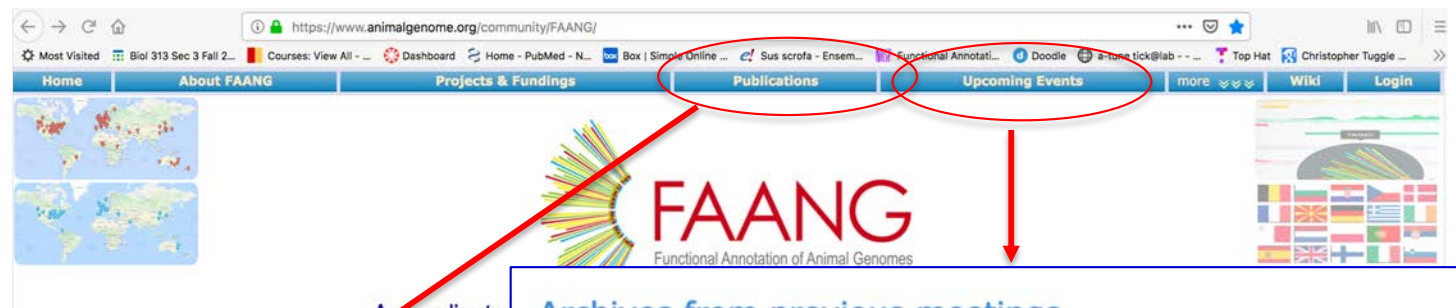
Kern, Zhou, etal., unpublished data

# ***Future of livestock functional genomics***

- **New** \$2.5 million Porcine FAANG project funded by NIFA-AFRI
  - *Headquartered at ISU*
- One of three funded (bovine, chicken) in 2018
- **Overall goal:** *catalog functional elements in the porcine genome for many biological states, create Chromatin state map*
- Aim 1 Adult tissues
  - extend FAANG pilot of 8 -> 25 tissues
- Aim 2 Fetal tissues
  - Correlate allele-specific expression and chromatin modification
  - Parent-of-origin effects (reciprocal crosses Meishan x WhiteCross)
- Aim 3 Immune System
  - Circulating Blood Cells- flow-sorted into functional types
  - *Single cell analysis of blood PMBC and immune tissues*

# Getting Involved: Resources

- FAANG website [www.faang.org](http://www.faang.org)
- FAANG Publications
- FAANG meetings (talk ppts available)



A screenshot of the FAANG Publications page. The page title is 'Functional Annotation of ANimal Genomes (FAANG) Project — A coordinated international action to accelerate Genome to Phenome'. The main heading is 'Publications'. Under 'Landmark Papers:', there is a review article by Elisabetta Giuffra et al. (2019) in Genomes (FAANG) and a GO-FAANG meeting paper by Tuggle et al. (2016) in Anim Genet. There is also a white paper link for 'The FAANG Consortium (2015), "Coordinated international action to accelerate Genome to Phenome, The Functional Annotation of ANimal Genomes (FAANG) Project". Genome Biology 2015, 16:57.'

A screenshot of the FAANG Aims section. The heading is 'A coordinated international action to accelerate Genome to Phenome'. Below it, the text reads 'FAANG Aims:'. The aims listed are: 'Standardize core assays', 'Coordinate and facilitate data sharing', 'Establish an infrastructure for data integration and analysis', and 'Provide high quality functional annotations of genomes'. There is a 'Sign up here to...' link and a 'FAANG data portal' link at the bottom.

## Archives from previous meetings

- FAANG at PAG-XXVII (PAG 2019)  
**Dates:** Jan. 11, 2019  
**Location:** Town & Country, San Diego, CA  
**Archives:** Power point slides, PDF, etc.  
**Status:** Held as planned
- 7th International Symposium on Animal Functional Genomics & Functional Annotation of Animal Genomes Workshop 2018  
**Dates:** Nov. 12-15, 2018  
**Location:** Adelaide, Australia  
**Archives:** Proceedings (PDF)  
**Status:** Held as planned
- FAANG at PAG-XXVI (PAG 2018)  
**Dates:** Jan. 12, 2018  
**Location:** Town & Country, San Diego, CA  
**Archives:** Power point slides, PDF, etc.  
**Status:** Held as planned

# Getting Involved: Resources

- **FAANG data portal: <http://data.faang.org>**

**Access to FAANG data limited to members → agree to share data and not publish first on others data**



The screenshot shows the FAANG website homepage. At the top, there is a navigation bar with links for Home, About FAANG, Projects & Fundings, Publications, Upcoming Events, Wiki, and Login. Below the navigation bar is a world map showing global participation. The main heading is "FAANG Functional Annotation of Animal Genomes" with a logo consisting of colorful lines radiating from a central point. Below the heading is the tagline "A coordinated international action to accelerate genome to phenome". The page is divided into several sections: "FAANG Aims:" which lists four bullet points (Standardize core assays, Coordinate and facilitate data sharing, Establish an infrastructure for analysis, Provide high quality functional annotation); "Working groups" which lists four groups (Steering Committee, Animals, Samples, and Assays (ASA), Bioinformatics and Data Analysis (B&DA), Communication (COM), Metadata and Data Sharing (M&DS)); and a section for newsletters and methods. At the bottom, there is a link to the "FAANG data portal" which is circled in red. A red arrow points from the text box on the left to this link.

FAANG  
Functional Annotation of Animal Genomes

A coordinated international action to accelerate genome to phenome

**FAANG Aims:**

- Standardize core assays and experimental protocols
- Coordinate and facilitate data sharing
- Establish an infrastructure for analysis of these data
- Provide high quality functional annotation of animal genomes

**Working groups**

- Steering Committee
- Animals, Samples, and Assays (ASA)
- Bioinformatics and Data Analysis (B&DA)
- Communication (COM)
- Metadata and Data Sharing (M&DS)

September newsletter of FAANG Europe COST Action

Introduction to FAANG Methods (Available since Dec 4, 2017)

**FAANG data portal:**



# FAANG data publicly available (circa October 2018)

<http://data.faang.org>

Interested in submitting your  
molecular data to FAANG?

We want your data!

Plans to set standards for data to  
be included in meta-analyses.

RNAseq (ENCODE):

- stranded
- >30M aligned reads

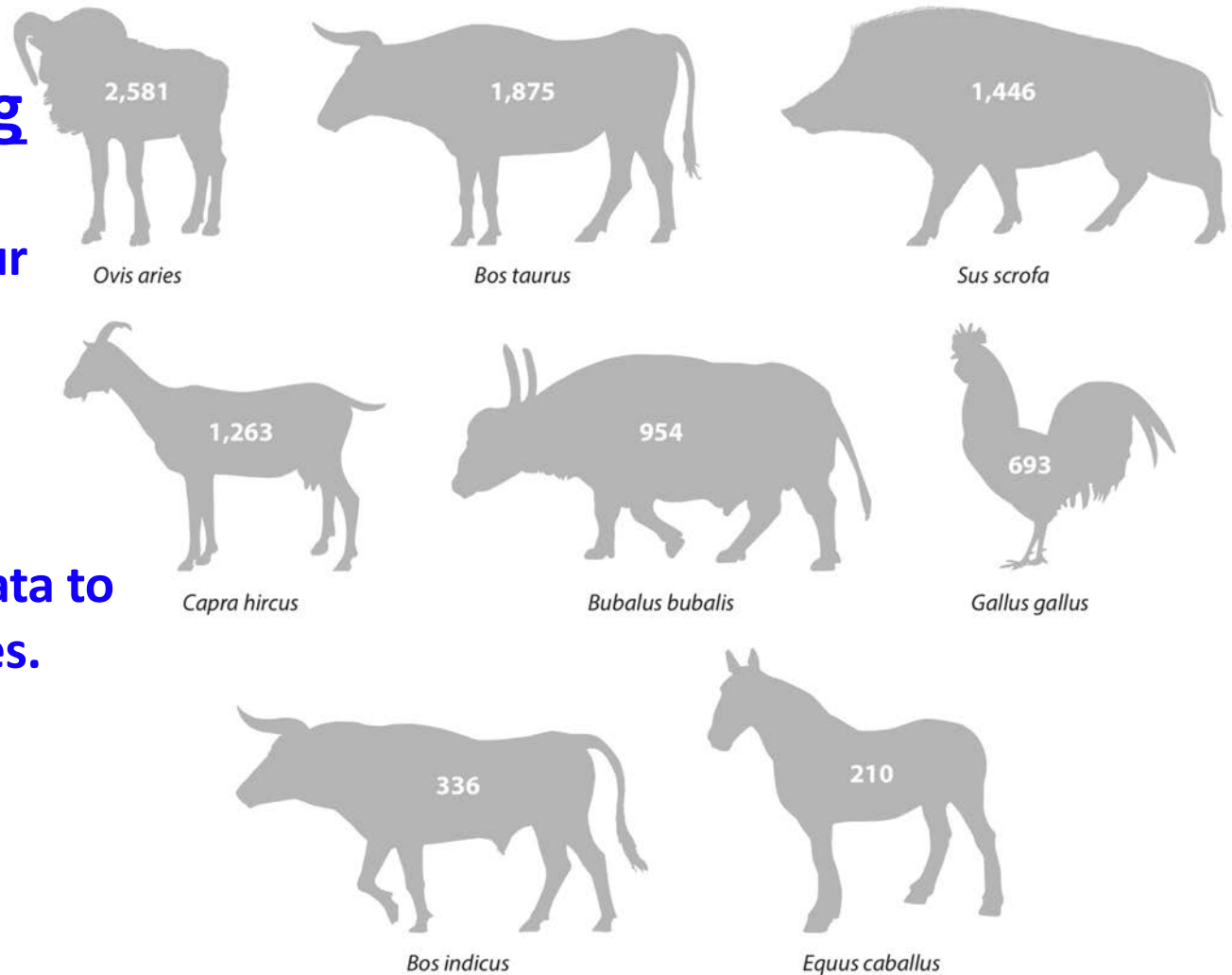


Figure 4

Registered specimens in the Functional Annotation of Animal Genomes (FAANG) data portal by species (as of October 2018) data are available from <http://data.faang.org/home>.

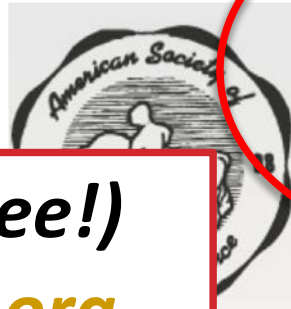
# *Conclusions and Outlook*

- > \$40 Million for FAANG world-wide and more to come
  - ~\$7.5M from USDA-NIFA-AFRI (**\$6 M started in 2018**)
  - ~\$20M from EC H2020 (to start in 2019)
  - ~\$5M from Australia
  - ~\$7M from Canada
- Next five-ten years will see an explosion in functional data in the cattle, pig, chicken, sheep and horse genomes
- ***Realizing a Predictive Biology- increase usefulness to society and industry***

# *Institutions supporting FAANG*



**GenomeCanada**



**Join FAANG (it's free!)**

**Visit: [www.faang.org](http://www.faang.org)**

**Email us:**

**[faang@iastate.org](mailto:faang@iastate.org)**



**IOWA STATE UNIVERSITY**  
OF SCIENCE AND TECHNOLOGY