*Functional **An**notation of **AN**imal **G**enomes (FAANG)-- Global FAANG Consortium*

Organic growth of FAANG during 2014-2019
Current FAANG contributors= 432

2014 membership- GB authors

March 2019 membership

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

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

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

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

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

What is functional annotation?

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

AGGGCCGTTGGGCTCCTCTGATTGGACGTGGCTGCGCCTGGCTGGCCAATGGGGAGATGGCGGCAGGTTCCCGGAAGTGGTGGCGCGGTCAGCTGCAGCTTGCATAACC
GCGGAGGGCTGGCGCCCAGTCGGCTGTGTTCGCCAACGCTATGAGTTCCTTCGAGGGCCAGATGGCGGAGTACCCAACTATCTCCATAGACCGTTTCGACCGGGAGAATC
TGAGGGCTCGCGCTTATTTCCTGTCCCACTGCCACAAGGGTGAGTGAGCGCGGCGCGCCGACCCCCTCCCGGGGACCAGGGCTGCGGCGGGTCTGGCCCCGCGGGGAG
GTAGCCCAGGGGCTGGAGAGGAGGAAGTTGGGGTGGGGGTCCTGCAGGAGGAGTGCAGGCTTTGGGAGGTCTGGGGAGAGACTAGCGACTGCATCCTGTCGCAGGC
GTGCTCTTCAGCCTGTTTGAGGCTTTATCGCCTCATACCTCCCCCCCCCCCCATCACACCCATCCCACTCCCGCTTACTCACTCACCATCACCATGTTTGATCAGGCACAGCTA
TTCGGTTTTTTTGGAAGCTCTCCCAGGGCCCAAT...

This is not English!

What is important information and what isn't??

GTGGTTTTTAACTTTTTGGGGAAGCTGACGTCCTTGGATGTGCTGTGGGCTGGCTTGTGTCCTCTCCCTACCCCCATTCATCTTTTGAATCCCTAATCTCCGATGTGGCTGTA
TTTGGAGAAAGGGCTTTTAGGAGGTGATGGAGGTTAAATGAGGTCATAAGGGTGGAGTCCTAATTTTATAGGAGTGGTGGCCTTCTCTACACACACACACACACACACAC
ACACACACACACCACACTGAGGAATGCCGTGTAGGCACGCGGCATGAAAGTGGCTGCTGGTAAGGAGAGGCGTCCCCAGAAACCAGCTTTGGTACCTTGATCTTGGACTT
GAGCTTCAGAACGTTAGAAATACATTTCTGGGTGTTTGCACTGTGGCACAGTGGGTTAAAAACCTGACTGCAGTGGTCAGATCTCCTGTAGCCGTGTGGGTTCGTCC
CTGGCCCAGTGCAGTGGGTTAAGGATCCAGCCTTGCCACAGCTGTGGTGTAGGTTGCAGCTGCAGCTTGGAGTCAATCCCTGCCCTGGGAACTTCCATATGCTGCAGGGA
CAGCGGTAGGAAAAAACAAATAAACAAATAAATAAATAAATAAGGAGTTCCCATTGTGGTACAACAGGGATGAATCCAACTAGGAACCATGAGGTTGCAGGTTCGATCCC
TGGCCTCGCTCAGTGGGTTAAGGATCCGACATTGCCATGAGCTGTGGTGTGGGTCGAAGATGAGGCTCAGATCTGGTGTTGCTATGGCTGTGGTGTAGGCTGGCAGCTG

FAANG

Iowa State University
OF SCIENCE AND TECHNOLOGY

Annotation can help….

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

Why is functional annotation important?

Is trait-associated genetic variation located (<u>enriched</u>) in genes?

Yes!

GWAS SNPs significantly enriched in protein-coding regions

But also <u>enriched</u> in regions <u>flanking/outside</u> known coding regions!
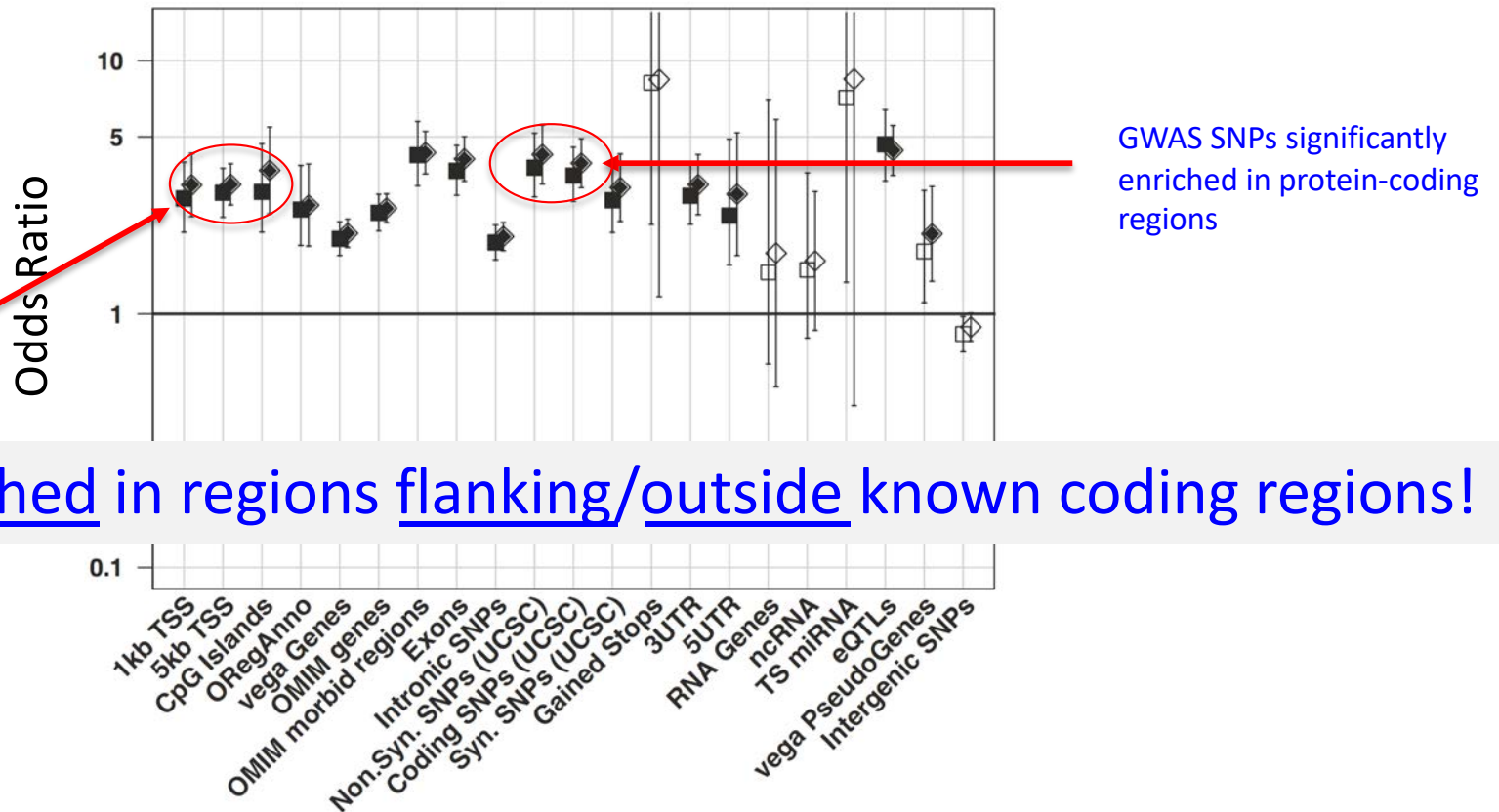
□ Permutations | 95% CI
◇ Random Sampling   • Significant Odds Ratios

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

Fat yield **Milk yield** **protein yield**
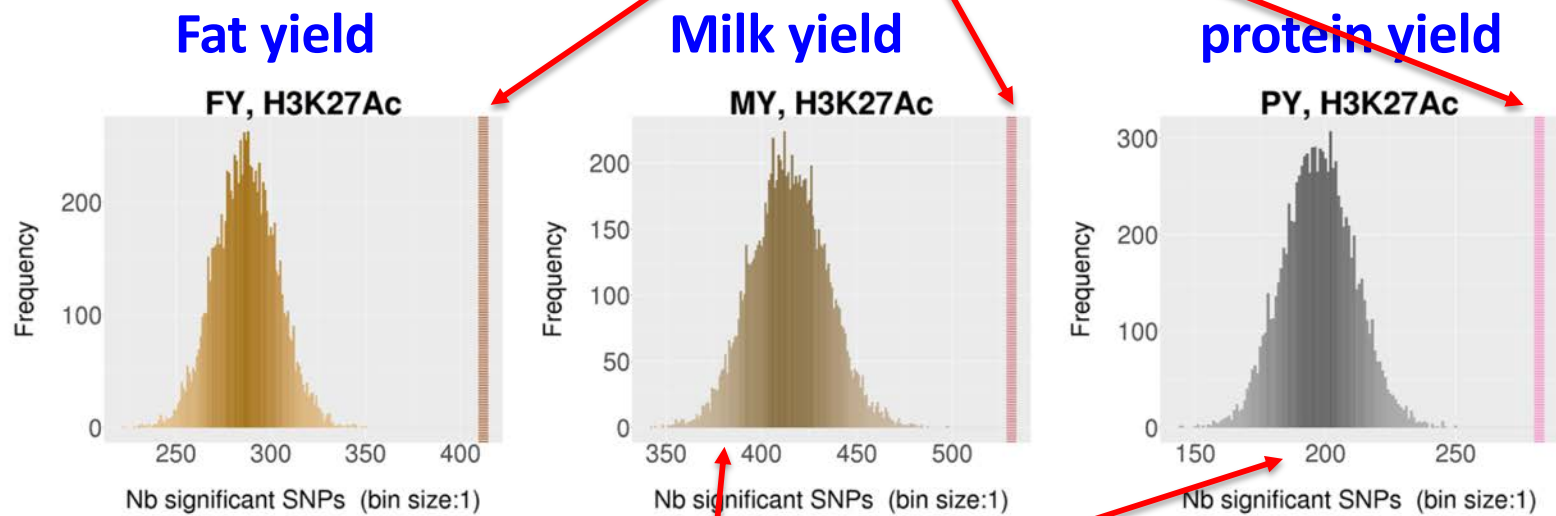
Much fewer sig SNPs in random sampling

FAANG
Functional Annotation of Animal Genomes

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- <u>which is the easiest parts to find and annotate</u>...
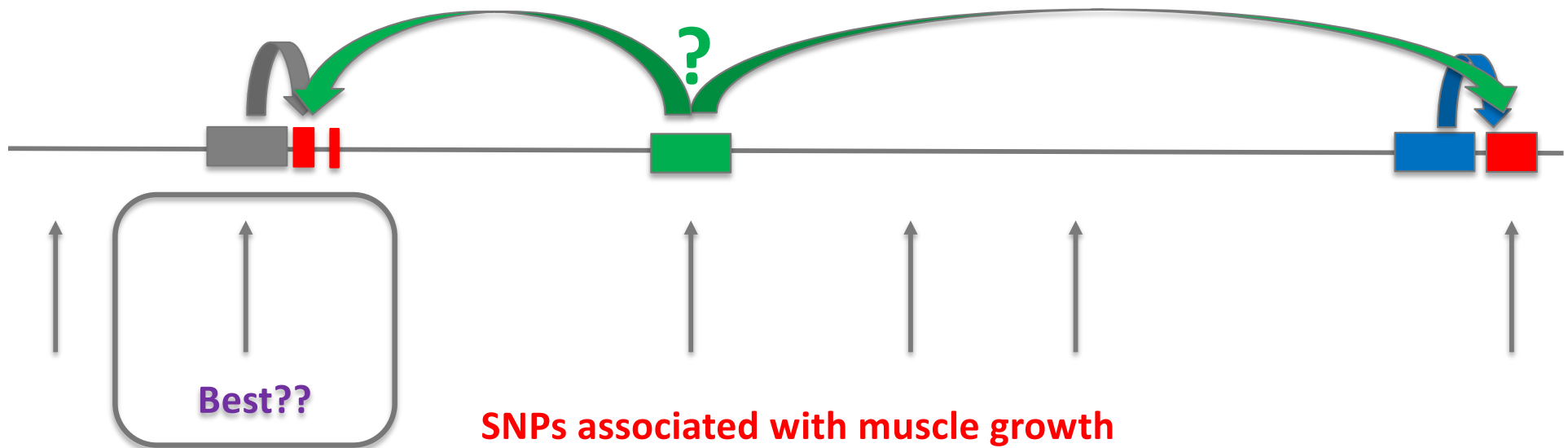
How do we figure out what parts are important then?

Protein-coding genes

Regulatory region- "ON" in skeletal muscle

Regulatory region- "ON" in mammary gland

Regulation- ???

Best??

SNPs associated with muscle growth

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

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 <u>all</u> 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?

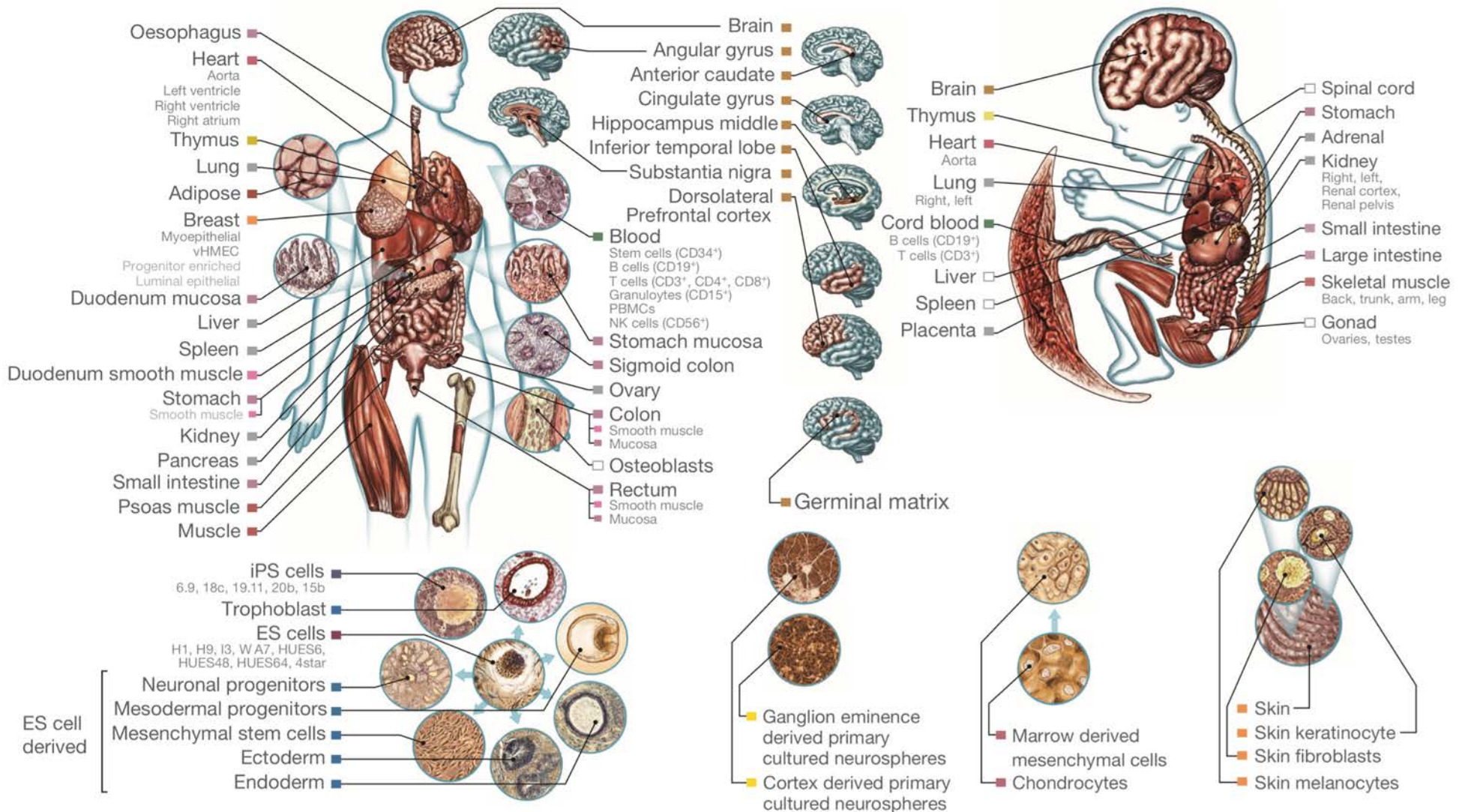b) what part controls this expression (specific tissue, response to infection, etc):
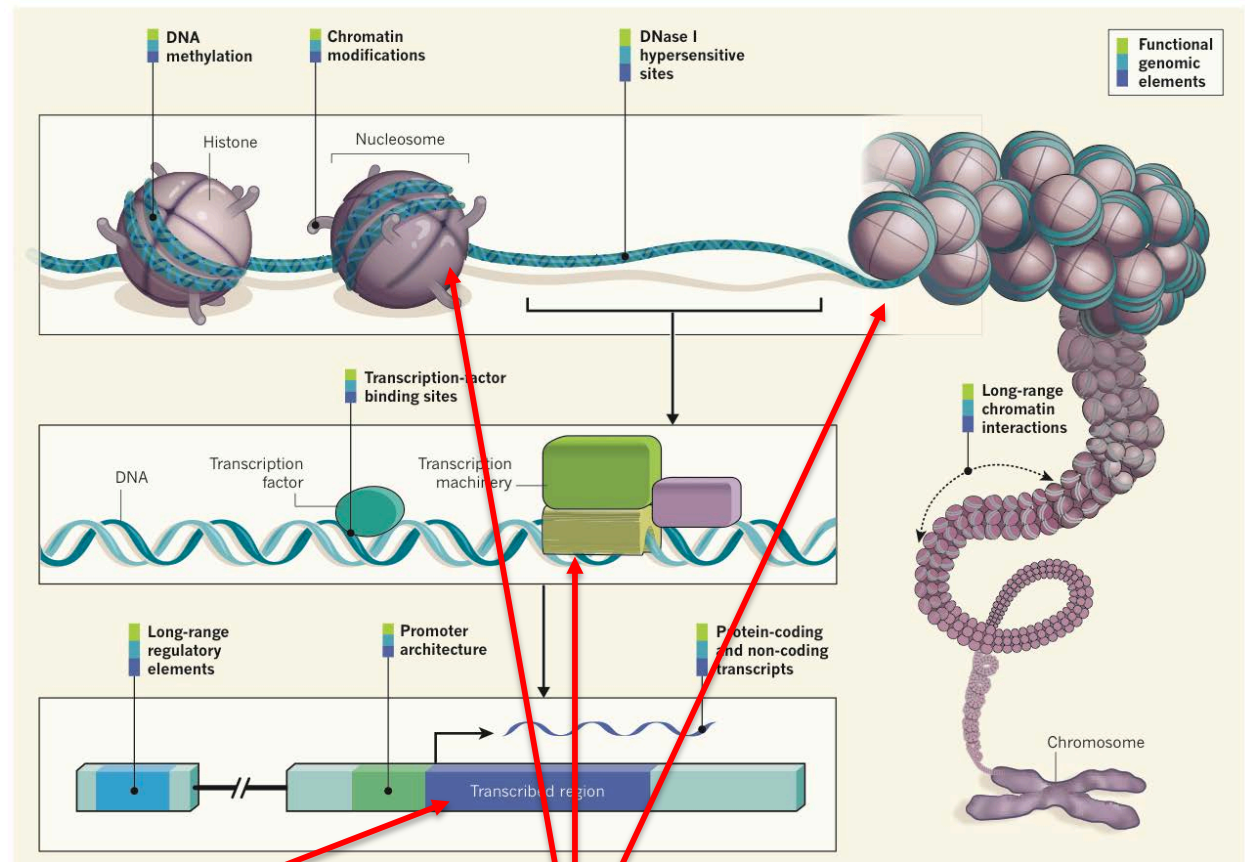
RNA levels and location

Biochemical assays of Chromatin structure

Roadmap Epigenome:
Extending ENCODE to tissues → many biological states
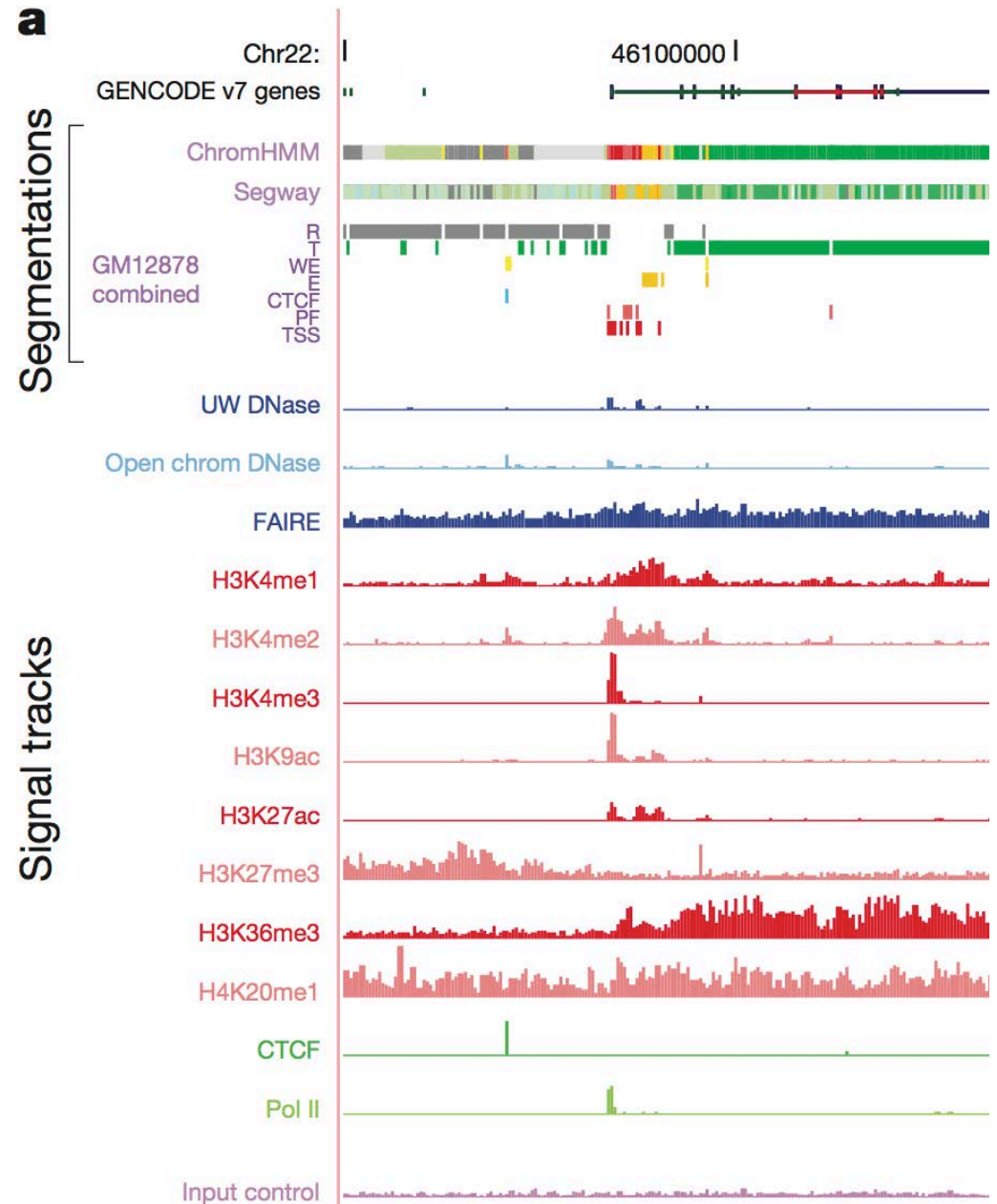
More biological states increases power to identify specific functional regions

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

FAANG
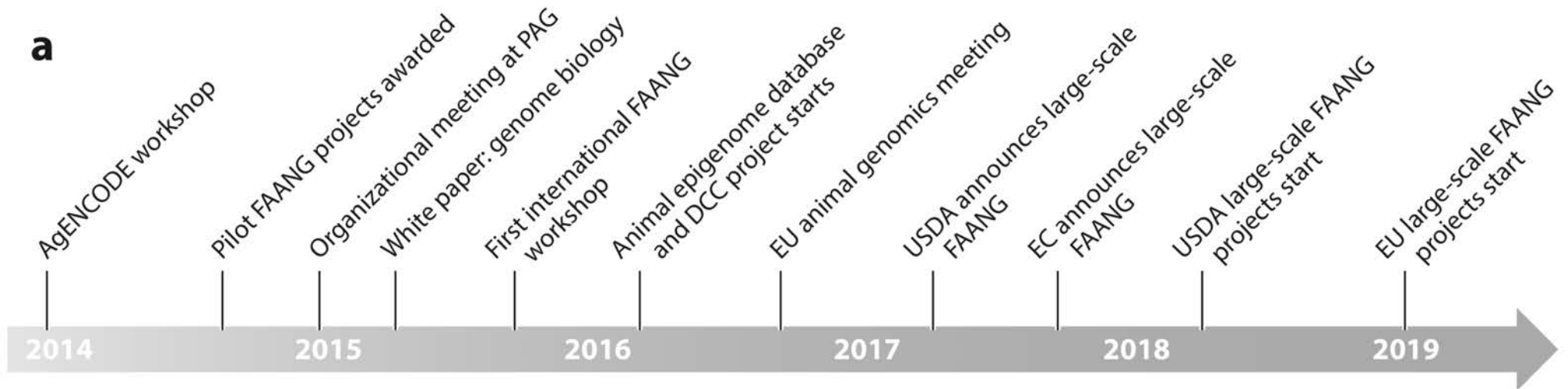Functional Annotation of Animal Genomes

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

a

AgENCODE workshop — Pilot FAANG projects awarded — Organizational meeting at PAG — White paper: genome biology — First international FAANG workshop — Animal epigenome database and DCC project starts — EU animal genomics meeting — USDA announces large-scale FAANG — EC announces large-scale FAANG — USDA large-scale FAANG projects start — EU large-scale FAANG projects start

2014   2015   2016   2017   2018   2019

FAANG
Functional Annotation of Animal Genomes

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

FAANG
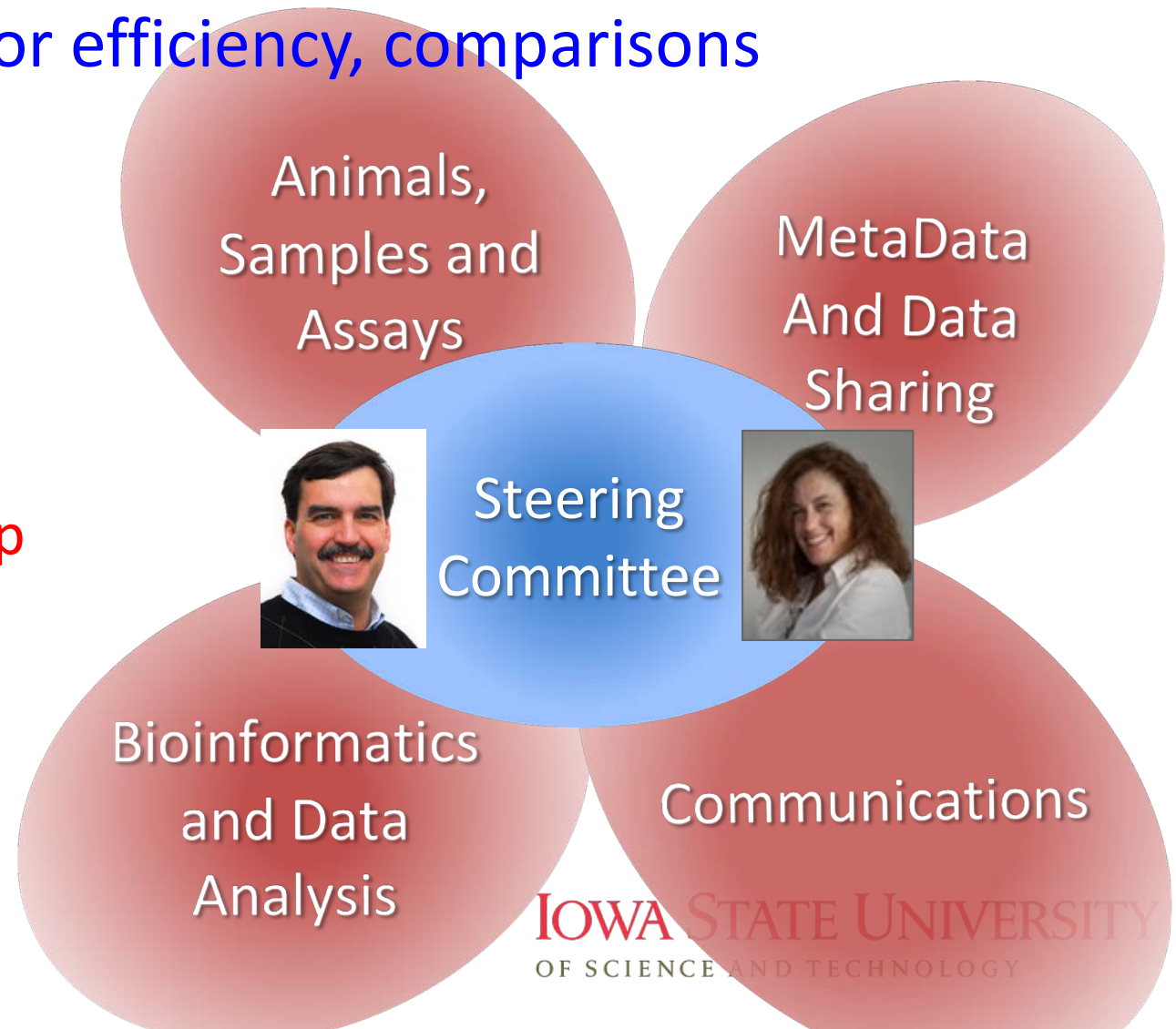Functional Annotation of Animal Genomes

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!

Animals, Samples and Assays

MetaData And Data Sharing

Steering Committee

Bioinformatics and Data Analysis

Communications

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

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.

Elisabetta Giuffra,[1] Christopher K. Tuggle,[2] and the FAANG Consortium*

[1]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
[2]Department of Animal Science, Iowa State University, Ames, Iowa 50011, USA; email: cktuggle@iastate.edu
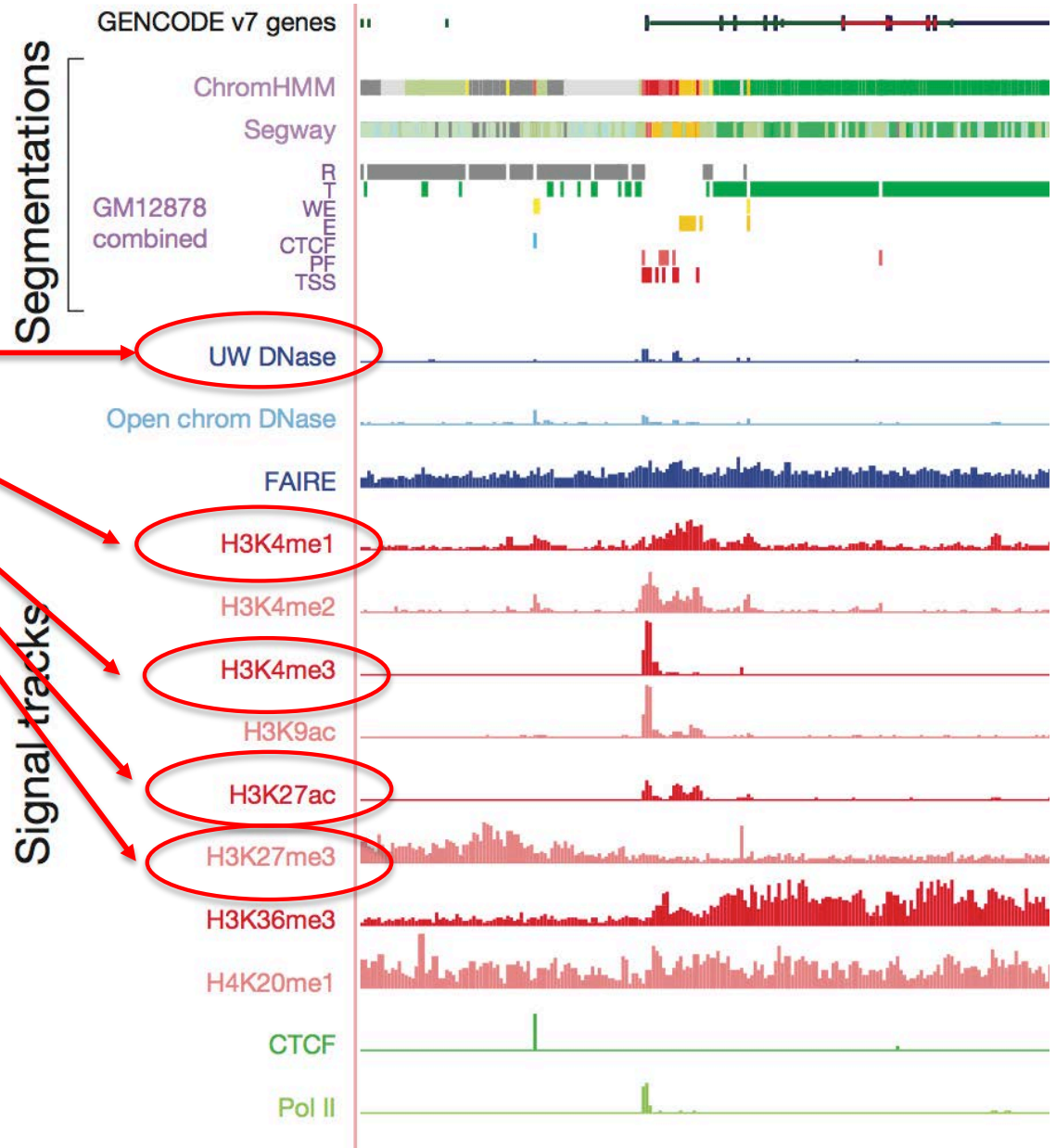
Main current activities

- **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**

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Minimum information for finding functional regions in livestock genomes

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-pilot	FAANG-rel.
pig, cattle, goat, chicken	INRA	France	Sylvain Foissac talk today	
pig, cattle, chicken	UC-Davis	US	X	x
horse	UC-Davis & Nebraska Univ.	US	Jessica Peterson talk today	
cattle	Leibniz Inst. for Farm Anim. Biology	Germany	X	x
Sheep	Int. Sheep Genome Consortium & others	Australia & US	Brenda Murdoch talk today	
			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

	Chicken		Pig		Cow
RNA-seq	Data Analysis		Data Analysis		Data Analysis
DNase/ATAC-seq	Data Analysis		Data Analysis		Data Analysis
H3K4me3	Data Analysis		Data Analysis	Chromatin Sheared	Data Analysis
H3K27me3	Data Analysis		Data Analysis	Chromatin Sheared	Data Analysis
H3K4me1	Data Analysis	Chromatin Sheared	Data Analysis	Chromatin Sheared	Data Analysis
H3K27ac	Data Analysis	Chromatin Sheared	Data Analysis	Chromatin Sheared	Data Analysis
CTCF	Data Analysis	Chromatin Sheared	Data Analysis	Chromatin Sheared	Data Analysis
RRBS-seq	Data Analysis		Data Analysis		Data Analysis
WTTS-seq	Data Analysis		Tissues Collected		Tissues Collected

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Example data for UCD pilot FAANG
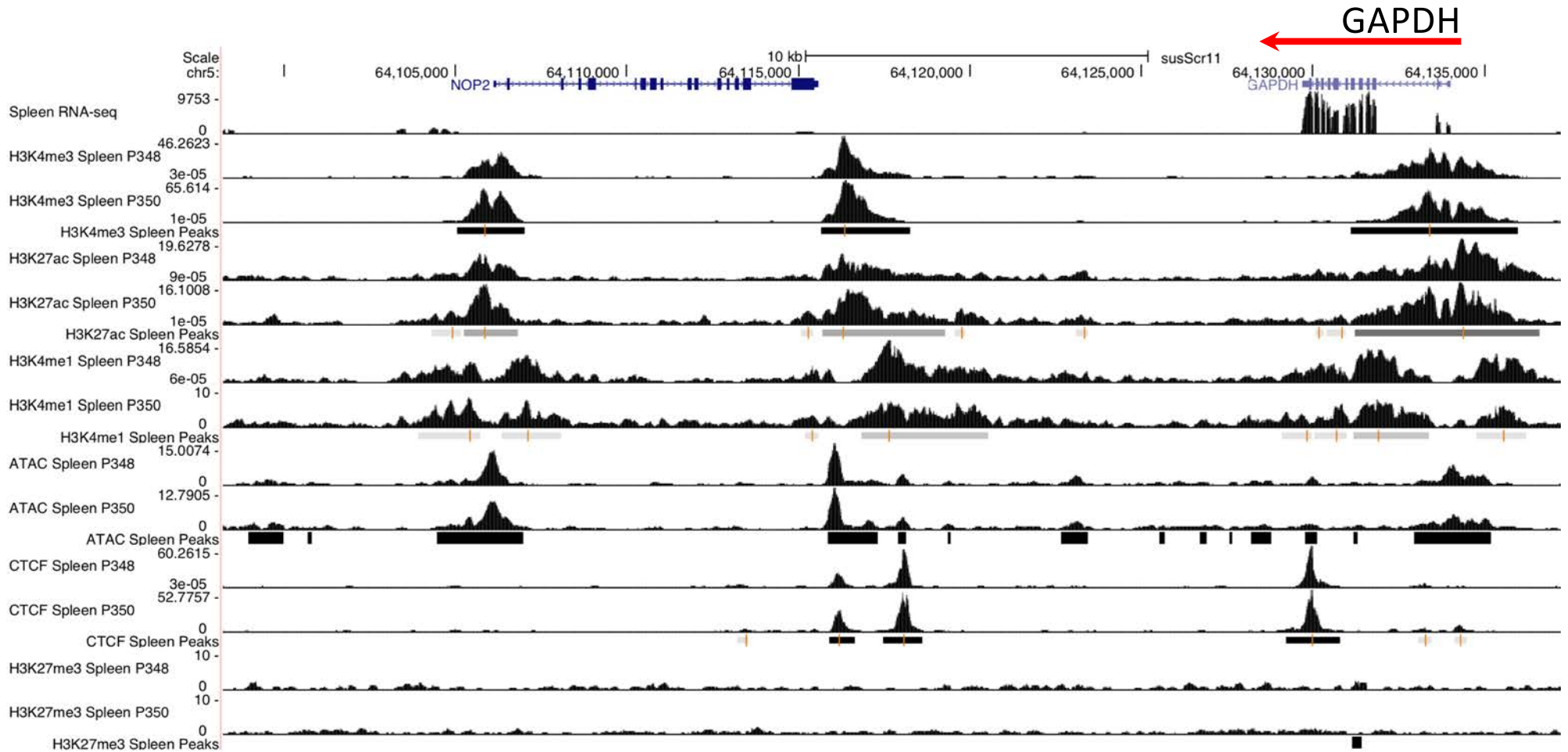
Kern, Zhou, etal., unpublished data

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-original 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*

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Getting Involved: Resources

- **FAANG website www.faang.org**
  - **FAANG Publications**
  - **FAANG meetings (talk ppts available)**

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

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

2,581 — Ovis aries

1,875 — Bos taurus

1,446 — Sus scrofa

1,263 — Capra hircus

954 — Bubalus bubalis

693 — Gallus gallus

336 — Bos indicus

210 — Equus caballus

FAANG
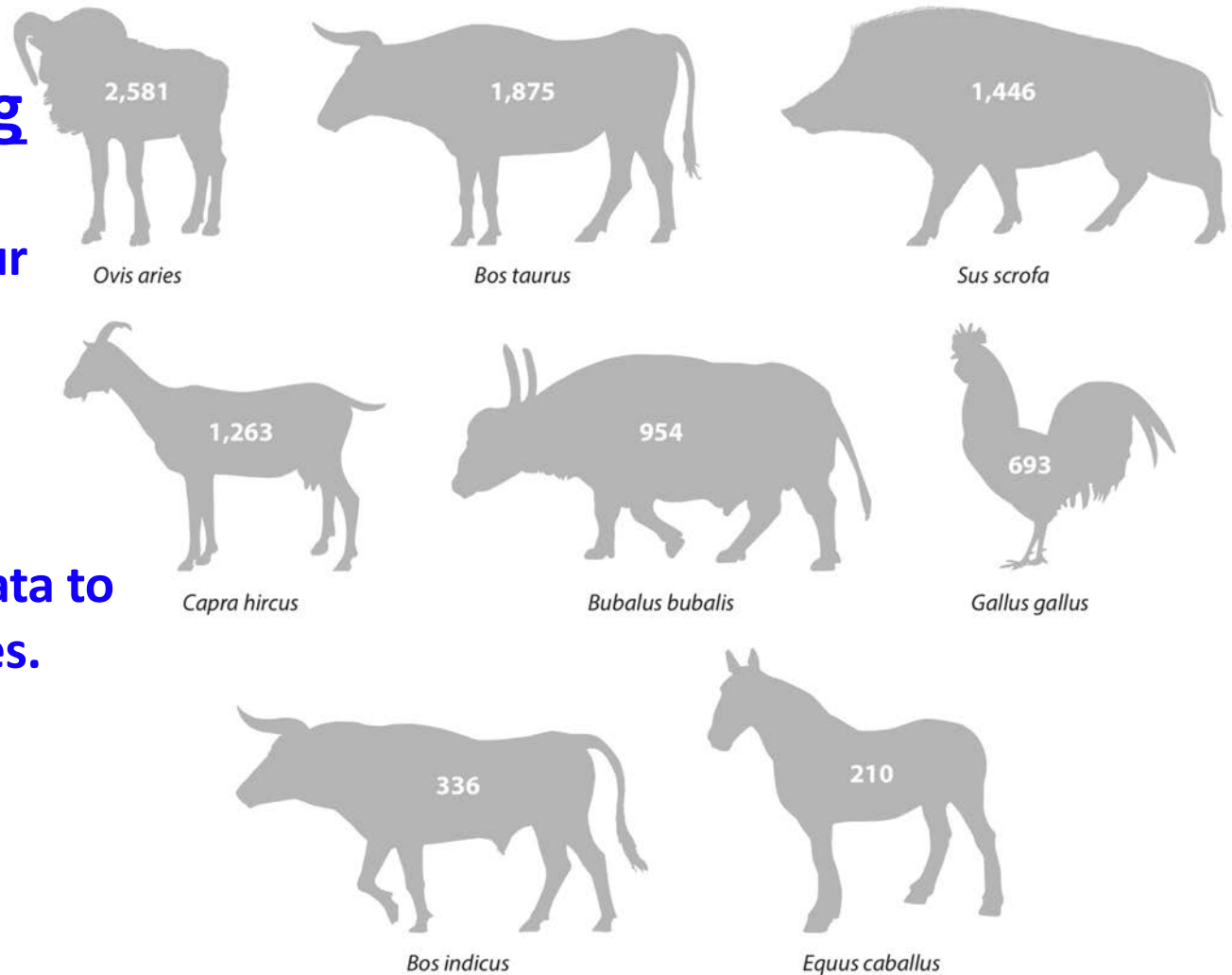Functional Annotation of Animal Genomes

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.**
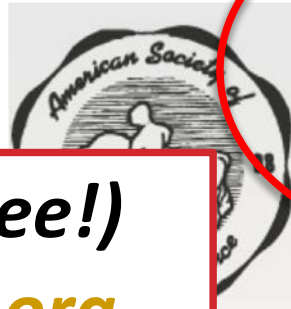
OF SCIENCE AND TECHNOLOGY

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*

FAANG
Functional Annotation of Animal Genomes

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Institutions supporting FAANG

Join FAANG (it's free!)
Visit: www.faang.org

Email us:
faang@iastate.org