

Farm Animal Genome Annotation: Current Status and Knowledge Gaps

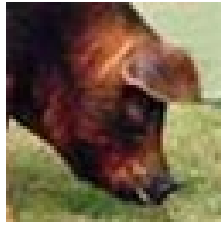
Fiona McCarthy
University of Arizona

State of Knowledge and Research Needs Regarding Heritable Genomic Modification in Food Animals: A Workshop
27-28 February 2024
National Academies of Science, Engineering & Medicine

OVERVIEW

1. Current status of farm animal genomes
2. Genome sequencing to support gene editing
3. Genome annotation of farmed animal species
 - The FAANG Project
4. Functional annotation to support gene editing
5. Annotation requirements to assess gene editing

Current status of farm animal genomes: Mammals



Duroc reference genome (BUSCO analysis 97.7%)
+ 57 additional breed & cross genomes

Hereford reference genome (BUSCO analysis 97.6%)
+ 74 additional *Bos* breed & cross genomes

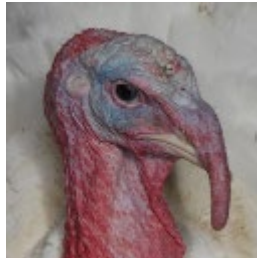


Rambouillet reference genome (BUSCO analysis 96.4%)
+ 57 breed & cross genomes

Current status of farm animal genomes: Poultry



Broiler reference genome (BUSCO analysis 98.6%)
+ 39 additional lines & cross genomes



Aviagen reference genome
Turkey_5.1 (BUSCO analysis 54.5%)

Current status of farm animal genomes: Aquaculture



Atlantic salmon reference (BUSCO analysis 54.5%)
+ 18 additional *Salmo* genomes



Channel catfish reference (BUSCO analysis 98.0%)
+ 4 additional *Ictaluridae* genomes

+ 22 different aquaculture species (including fish, prawns, oysters
which have been gene edited

- Growth
- Disease resistance
- Reproduction

Roy et al Front. Mar. Sci., 14 July 2022 REVIEW article
Sec. Marine Fisheries, Aquaculture and Living Resources
doi.org/10.3389/fmars.2022.924475

Genome Sequencing to support Gene Editing

- Genome editing is more precise when the genome sequence & annotation used for analysis is based on the same line or breed.
- Reference assemblies bias the analysis.
 - do not account for gene edit changes

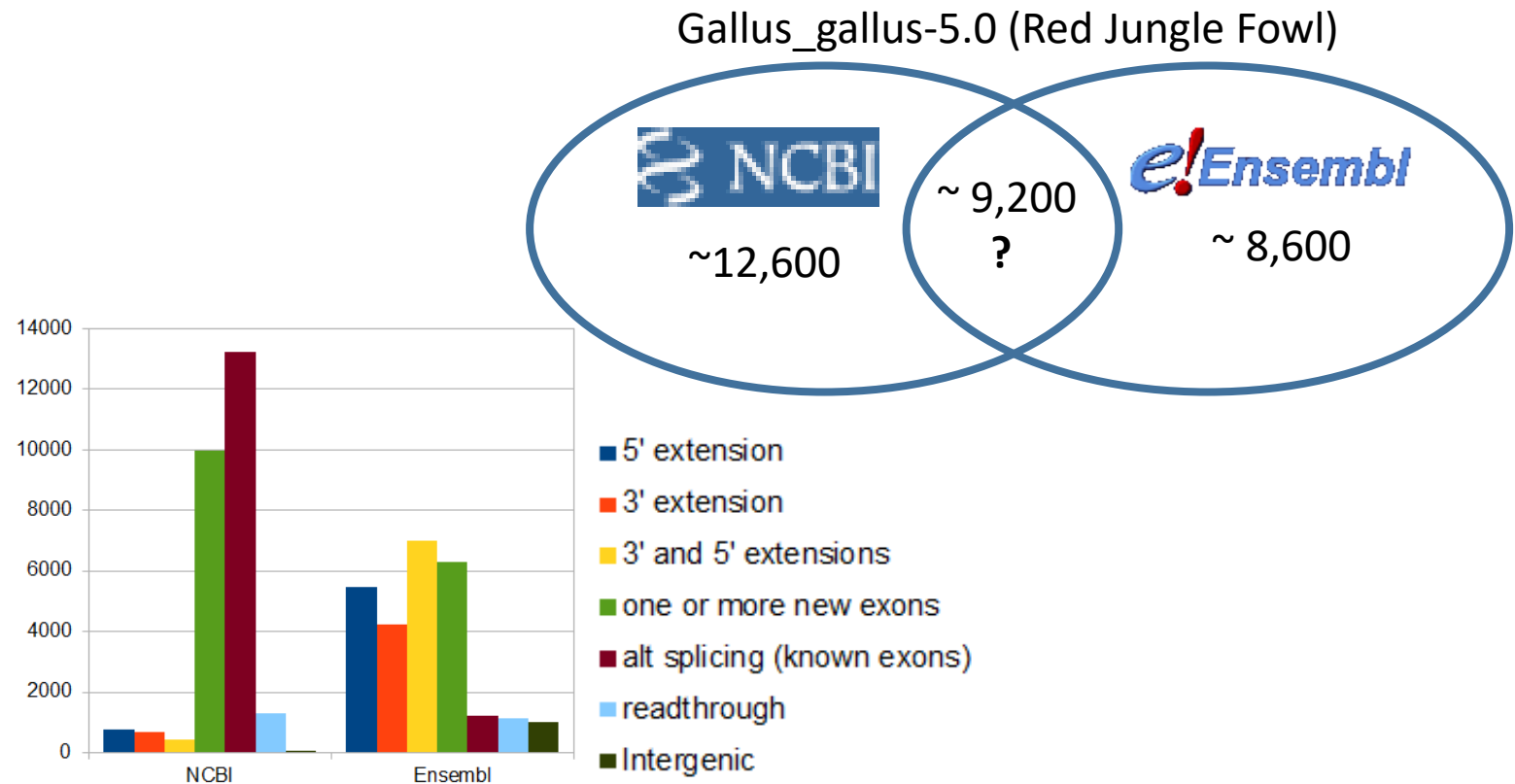
Genome Annotation is Crucial to Support Gene Editing

1. Clearly defined genomic elements are required to ensure defined gene edit sites.
2. Knowledge of regulatory elements supports prediction of when & where gene edits are likely to be expressed.
3. Ability to link changes in gene expression to changes in function
 - What does normal gene expression look like?
 - How do changes in gene expression impact physiological processes?
 - Use gene markers for stress physiology to measure welfare

There is no single annotation set for farmed animals.

- No MODs exist for livestock & aquaculture species
 - no central coordination for genome annotation
 - NCBI & Ensembl – provide two different annotation sets

➤ *for genome editing, the initial annotation starting set matters*



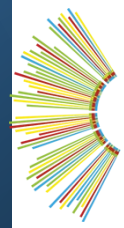
Genome Annotation

Two Complementary processes:

1. **Structural** – identifying functional elements within the genome sequence
 - genes & regulatory elements
2. **Functional** – linking functional information to genomic elements
 - gene product function, gene names, expression patterns

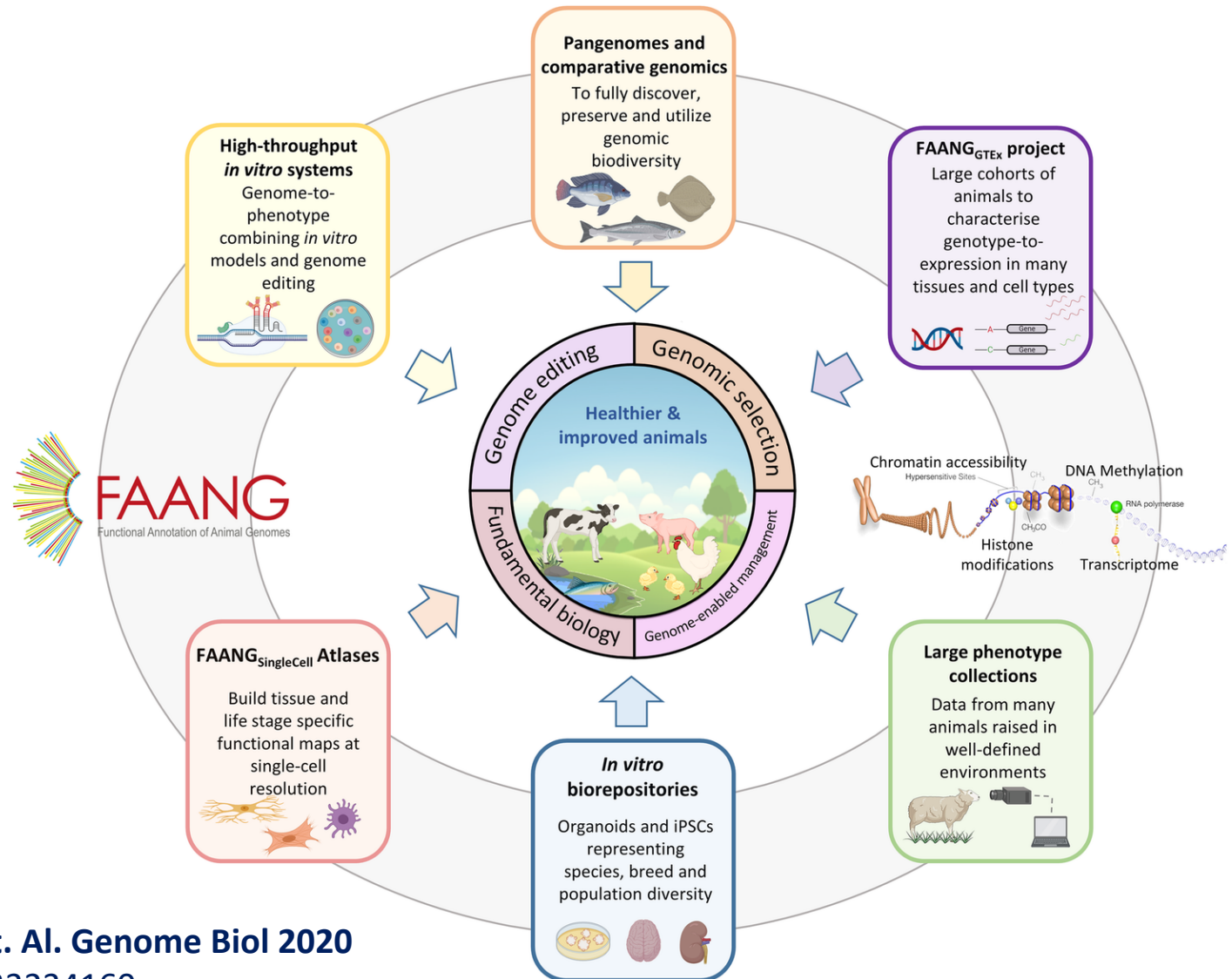


Structural Genome Annotation: The FAANG Project



FAANG
Functional Annotation of Animal Genomes

co-ordinated data generation to characterize expressed and regulatory genomic regions in farmed animal genomes

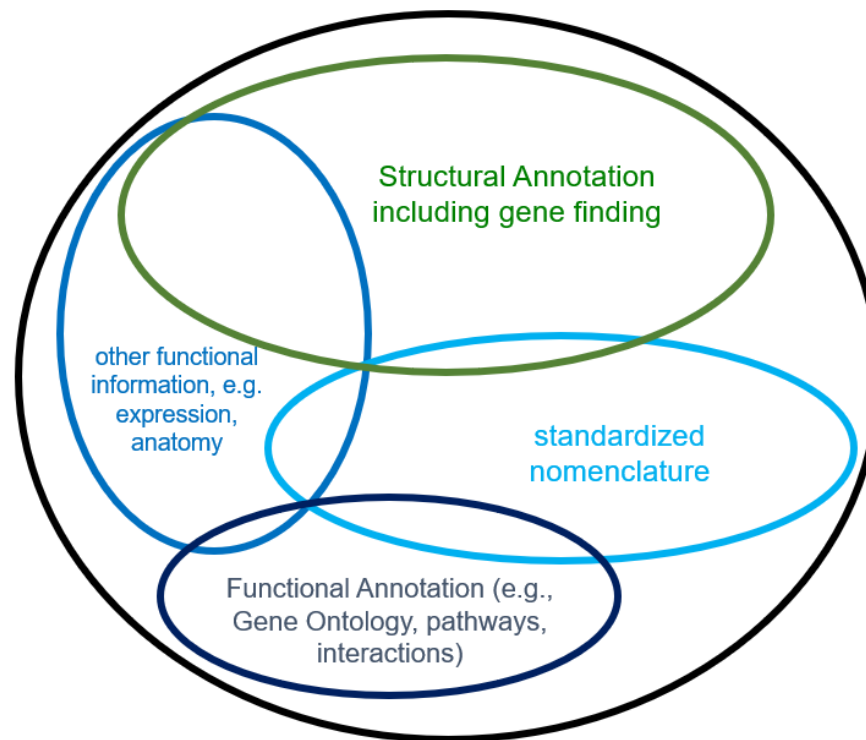


Structural Genome Annotation to support Gene Editing

- Many edits are made in the “noncoding” regions but without knowing regulatory elements, we can’t say for sure what is noncoding vs what is not annotated.
- Open chromatin and methylation data provides crucial information needed to assess when & where gene modifications are likely to be expressed.

Functional Genome Annotation

- Functional annotation is the (independent) process of adding functional information to genomic elements
 - Gene Ontology, Pathways, expression
 - mostly *ad hoc* in farmed animals
 - Standardized gene nomenclature used to integrate annotations



VGNC



Functional Annotation to support Gene Editing

- In nearly every case, "normal gene expression" is ***not*** known.
 - difficult to determine if editing perturbs gene expression
 - many groups are collecting expression data but there are gaps, and a lot of the data is collected at the mRNA level rather than protein level.
- It is not clear when a change in expression has a phenotypic effect
 - requires a shift from looking at individual gene expression to determining gene editing impacts on biological pathways and processes
 - biological pathways annotation is poor; pathways analysis tools are worse

Annotation Requirements to Assess Gene Editing

1. Structural Annotation

- identifying genes to ensure intergenic gene edits
- annotating regulatory elements to understand expression environment

2. Functional Annotation

- understand “normal” gene expression contributing to physiological processes
- identify functional changes due to gene editing & their implications at a phenotype/physiological level
- develop welfare indicators from stress physiology markers