



# Genome and Epigenome eNabled breeding in MOnogastrics

*Frédérique Pitel & Tatiana Zerjal (INRAE) & Sandrine Lagarrigue (Institut Agro)*

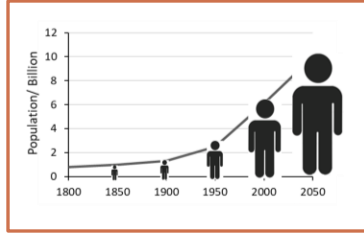
EAAP conference  
FAANG session

30 august 2023 – Lyon, France

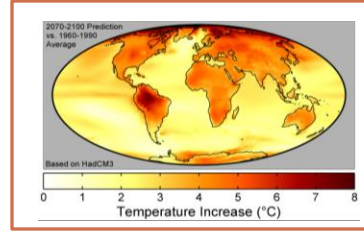


## Challenges

Human population growth



Environmental degradation



Food security



Changes in socio-cultural values



## Aim

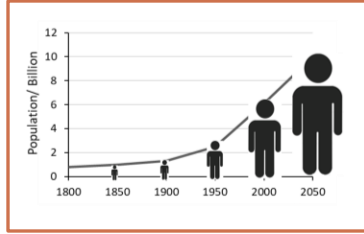
Rethinking animal breeding in pig and chicken



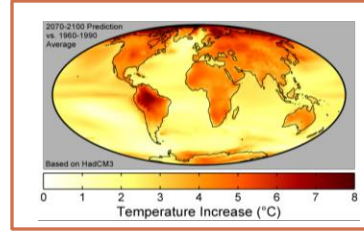
to move toward **a more sustainable model**, ensuring production but also promoting efficient resource use, animal resilience & welfare, and preserving genetic diversity.

## Challenges

Human population growth



Environmental degradation



Food security



Changes in socio-cultural values



## Aim

Rethinking animal breeding in pig and chicken

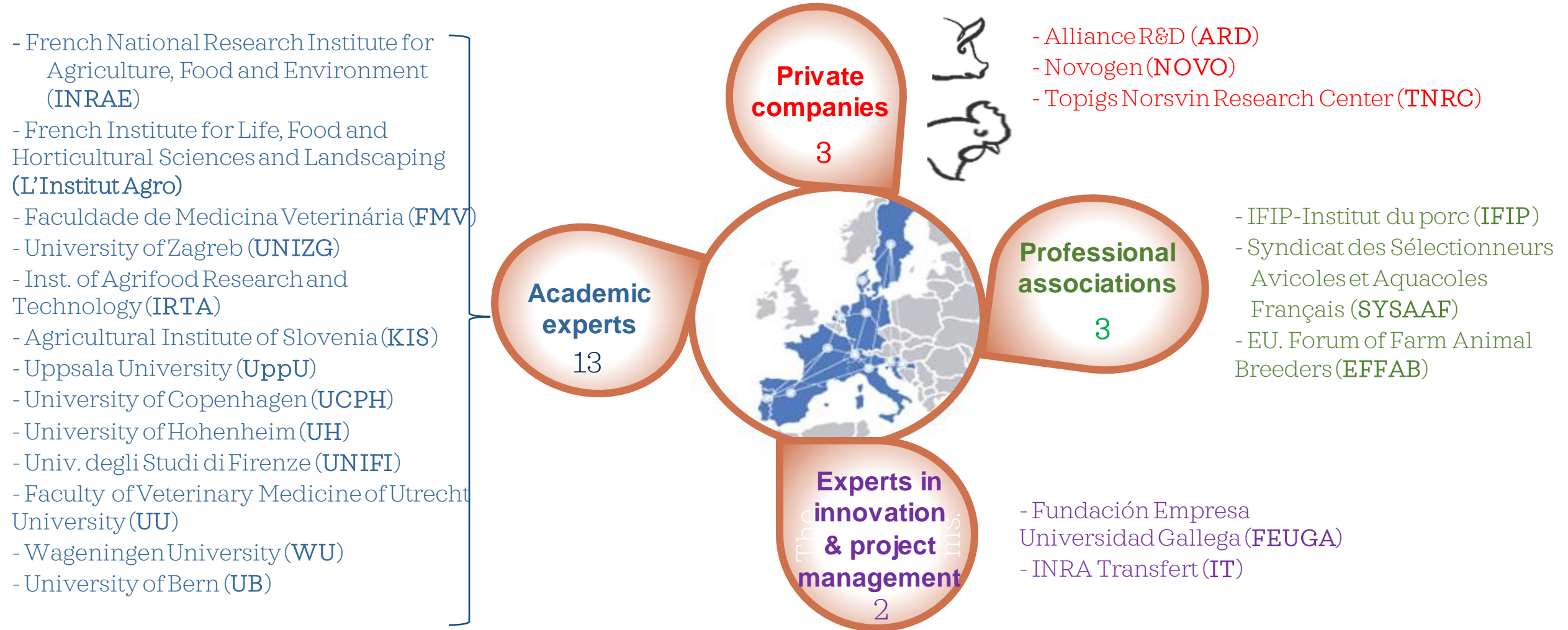


to move toward **a more sustainable model**, ensuring production but also promoting efficient resource use, animal resilience & welfare, and preserving genetic diversity.

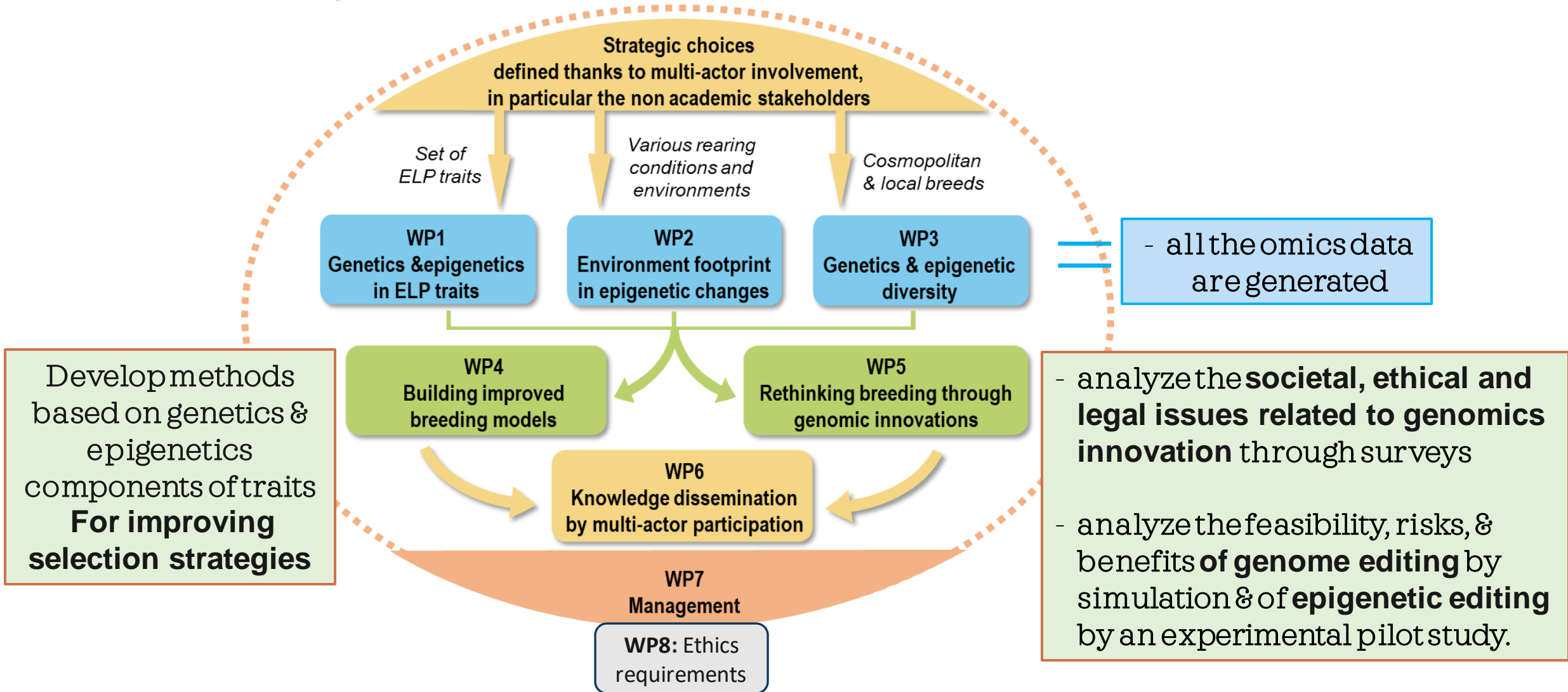
## Strategy

- Use **current “Omics” technologies** associated with **different phenotypes** related to production (quantity & quality), feed efficiency, productive longevity, fertility, resilience and welfare.
- Contribute to **improve knowledge on genome-to-phenome relationship** accounting for both **genetic and non-genetic mechanisms** controlling traits.

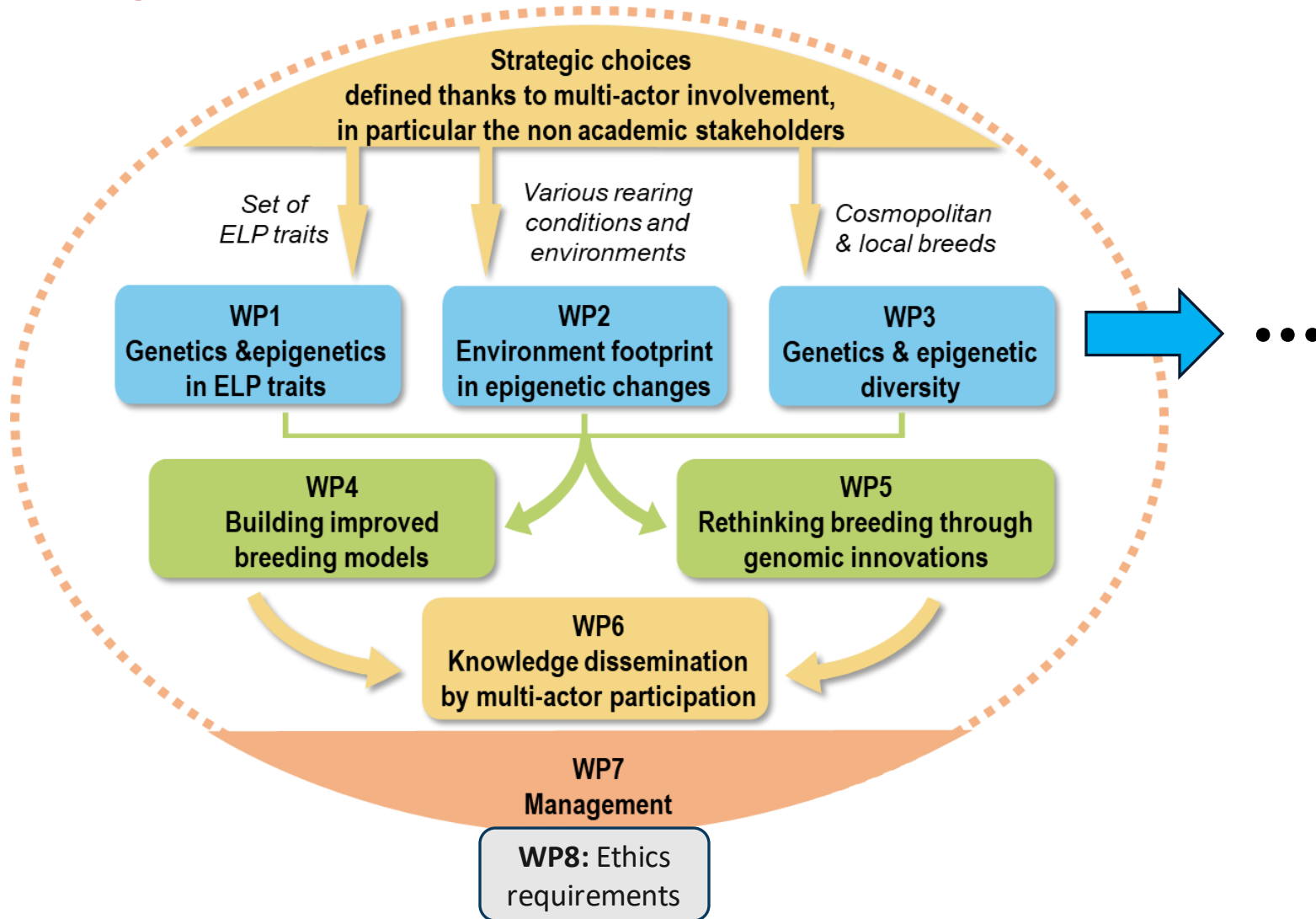
## Europe multi-actor project: 21 partners - 11 countries



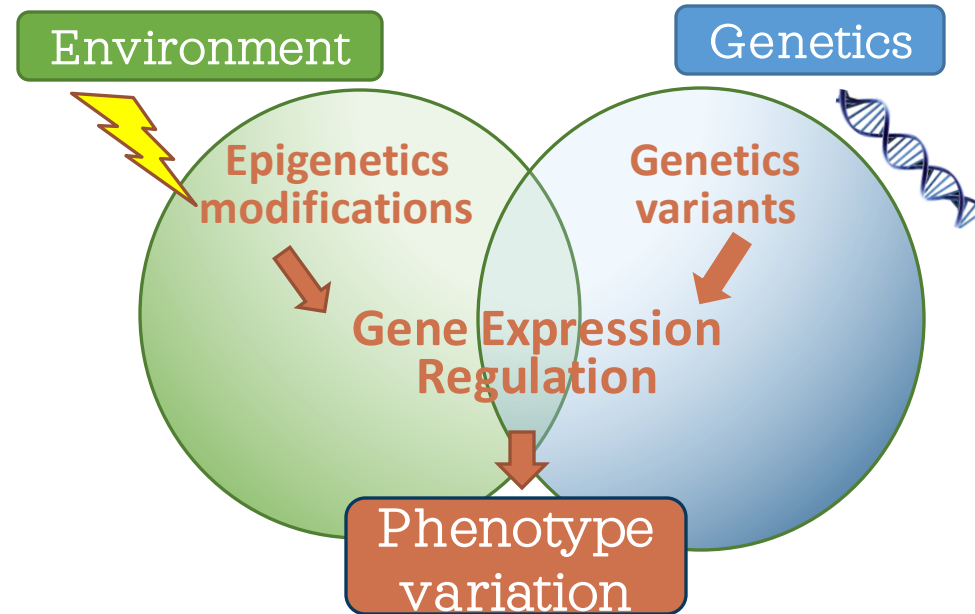
## 8 Work Packages (WP) with 5 scientific WPs



## 8 Work Packages (WP) with 5 scientific WPs



# GEroNIMO objectives

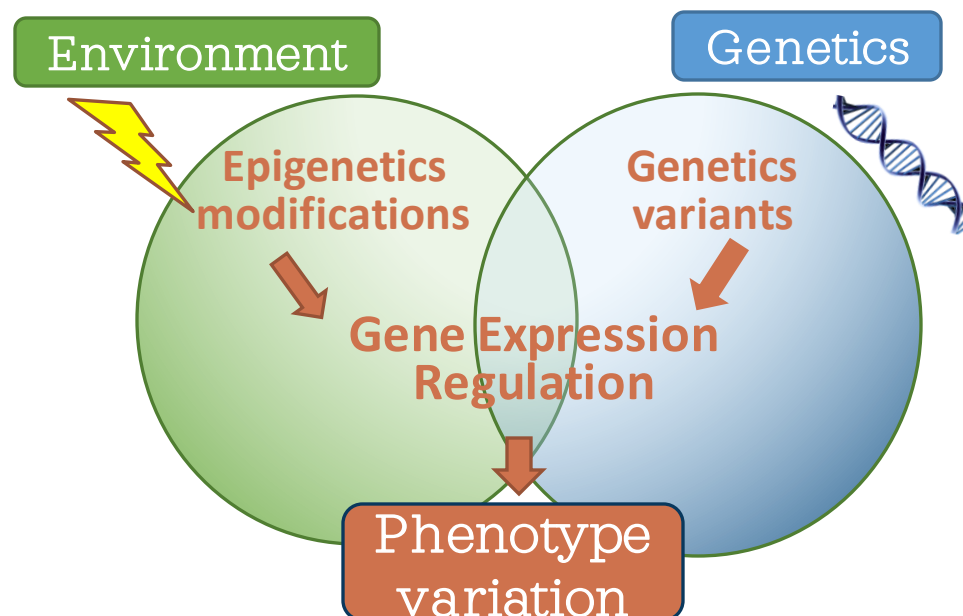


WP2

Identify biological mechanisms affecting trait variation considering **genetics & epigenetics components**

WP1

Large animal designs  
GWAS analysis



- Sperm quality
- Growth & Feed efficiency
- Laying rate, Egg quality & feed efficiency during long careers of production (over 90wk)

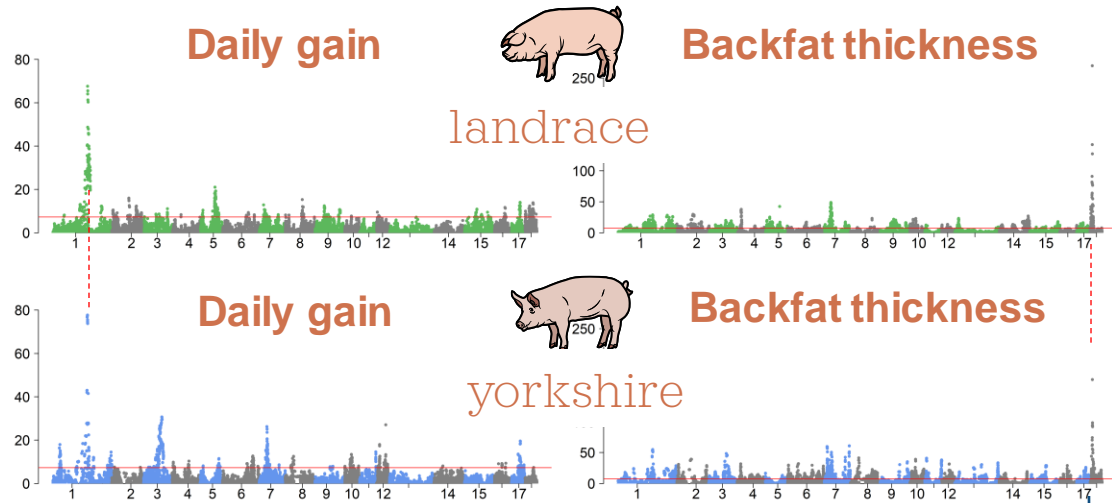




# A few results about the WP1

## WP 1.3

20,000 genotyped and phenotyped pigs/breed.  
Some QTL are shared by the two breeds



*SOCS6* = Leptin/insulin resistance

*PTN* = Adipogenesis  
*DGK1* = Fat deposition



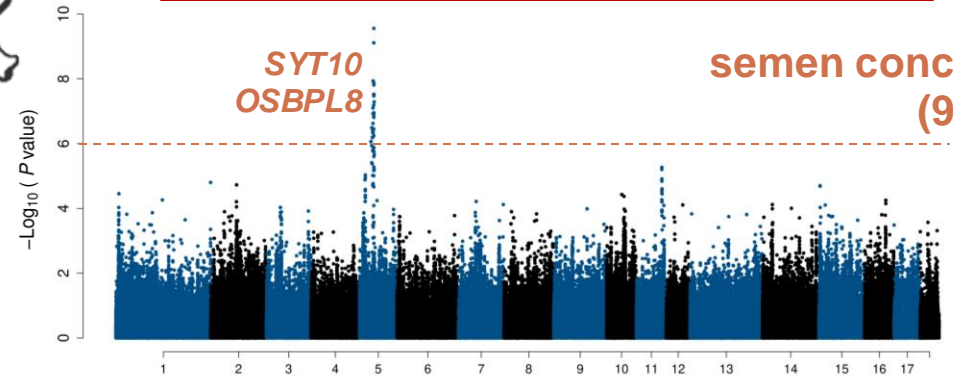
UNIVERSITY OF COPENHAGEN



Maria Luigi et al.

## WP 1.1

400 genotyped and phenotyped pigs



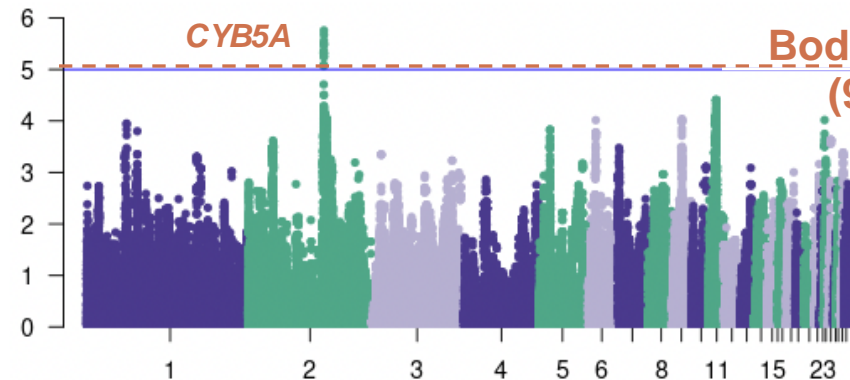
semen concentration (9 months)



Pedro Sá et al.

## WP 1.2

~1000 genotyped and phenotyped laying hens



Body fat reserve (90 weeks)



Alexandre Hubert et al.

Identify biological mechanisms affecting trait variation considering **genetics & epigenetics components**

WP2

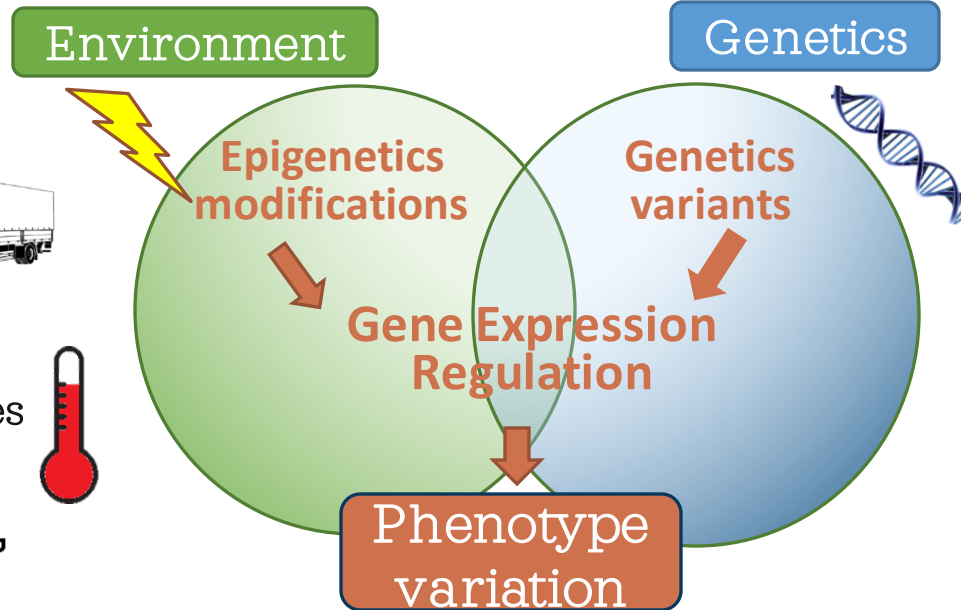
WP1

Smaller animal designs  
Compare # environments

Large animal designs  
GWAS analysis



- # Transport & rearing conditions 
- # Diets 
- # Ambient temperatures & seasons 
- # Ages or generations 



- Sperm quality
- Growth & Feed efficiency
- Laying rate, Egg quality & feed efficiency during long careers of production (over 90wk)



Map **G and epiG variations** in different chicken and pig **local breeds** to propose strategies for **optimizing the conservation of genetics and epigenetics diversities**

WP3



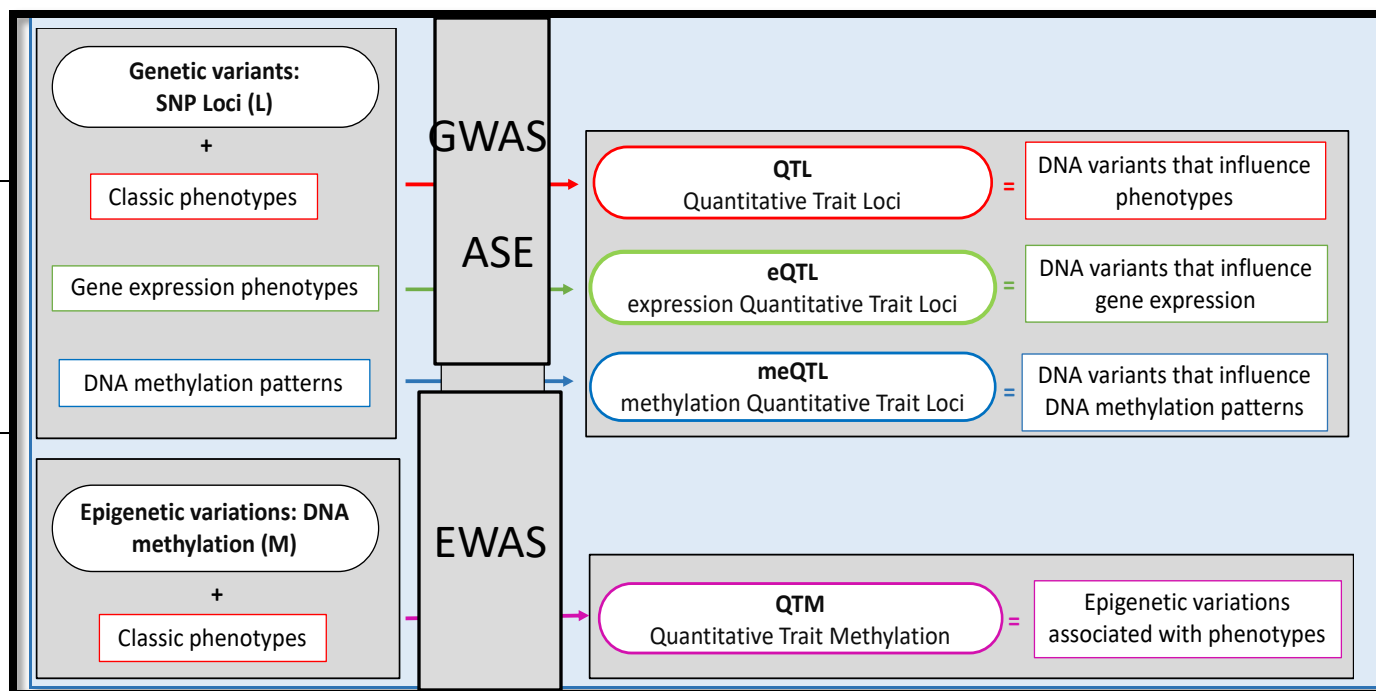
After 2 years, we have accumulated a large number of animals/samples/omics data

Genotypes and phenotypes from breeders or GEroNIMO	Transcriptomics data	epigenomics data with the associated Epigenotypes
> 90,000	2000	10,000

-PolyA RNA-seq +++  
 -Small RNA-seq +  
 -Ribo depleted RNA-seq +

*coming soon*



We perform Genetics & epigenetics analyses



to decipher the part of genetics & epigenetics involved in phenotype variation

# Pipelines used in GERO NIMO

- Pipelines from **nfCore** when they are available
  - “**nf-core/rnaseq**” for RNAseq analysis
  - “**nf-core/methylseq**» for bisulfite
  
- **In-house pipelines** combined with pipelines already approved by others
  - in-house pipeline (Uppsala University) for GBS-MeDIP analysis
  - **TensorQTL** (from broad Institute) for GWAS with gene expression
  - **phASER** (from Stephen Castel) for Allele Specific Expression

➤ <b>Genome assembly (.fa)</b>	➤ <b>Genome annotation (.gtf)</b>
 <p>Sscrofa11.1</p>	<ul style="list-style-type: none"> <li>• e! v107</li> </ul>
 <p>GRCg7b</p>	<ul style="list-style-type: none"> <li>• An improved version, = Combining gene models from different resources</li> </ul>



# We successively gathered 6 gene model resources

The 2  
reference  
genome  
annotations



Ncbi

1



Ensembl

2

including *Gene-Switch* *rnaseq* data

The 2 resources  
related to **FAANG**  
chicken  
**multi-tissue**  
RNA-seq data



Fr-AgENCODE

3

UC Davis

4

2 additional  
resources



INRAe

5



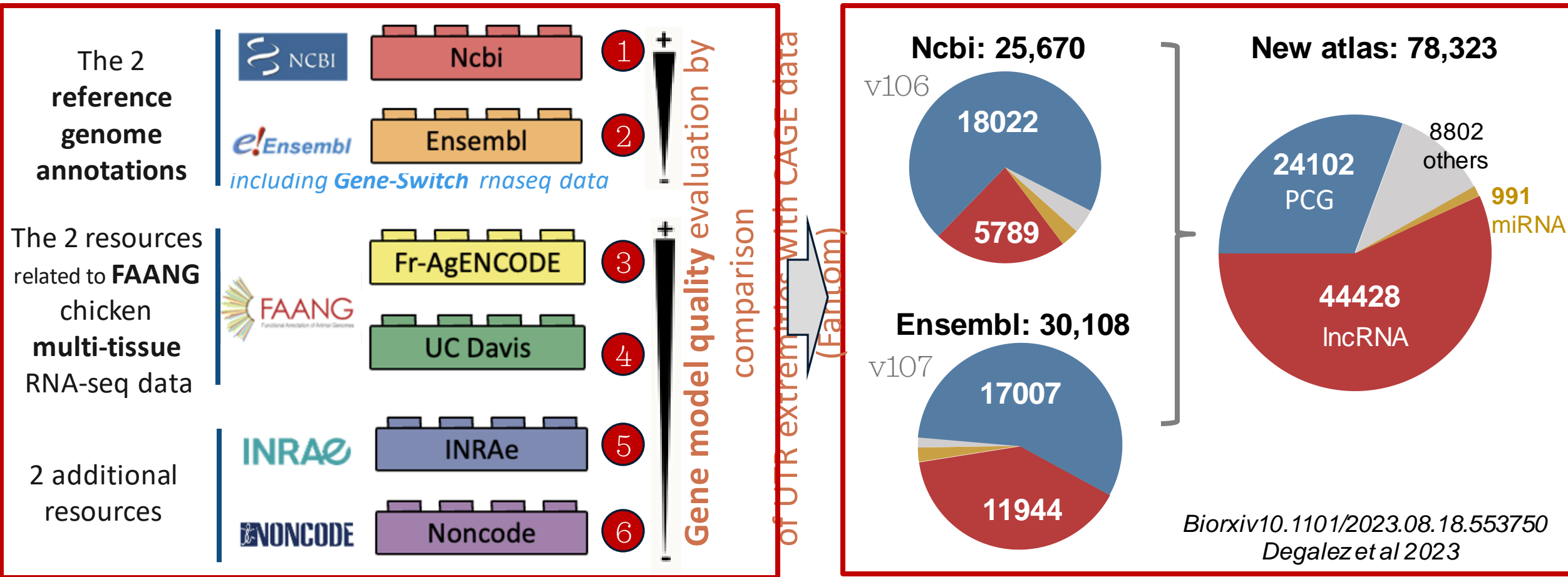
Noncode

6

Gene model quality evaluation by  
comparison  
of UTR extremities with CAGE data  
(Fantom)

# We successively gathered 6 gene model resources

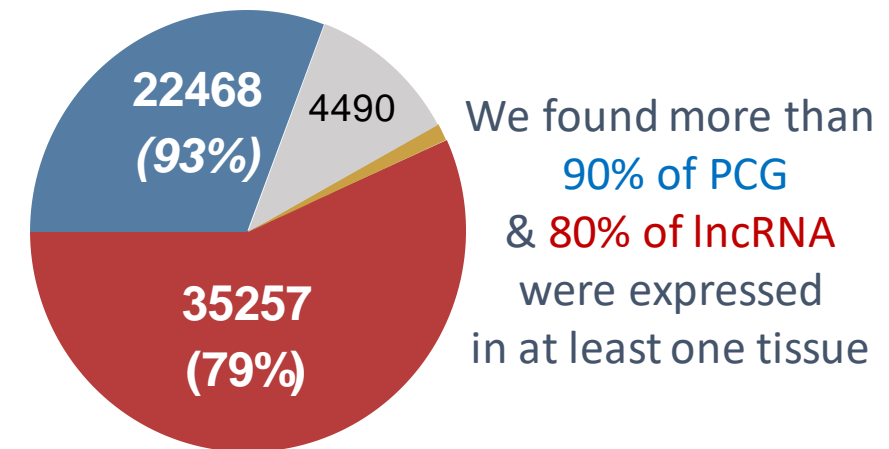
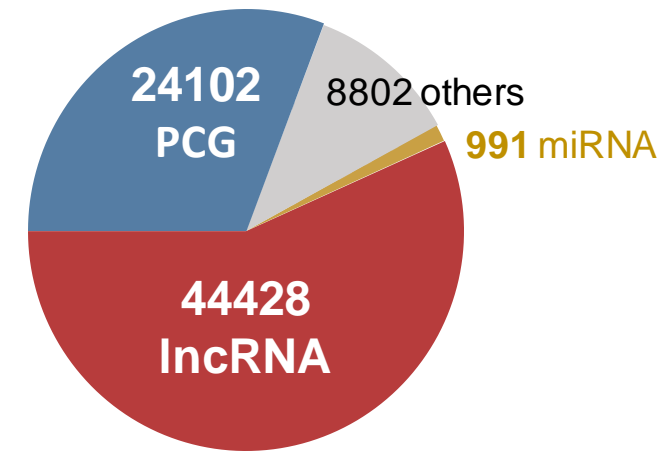
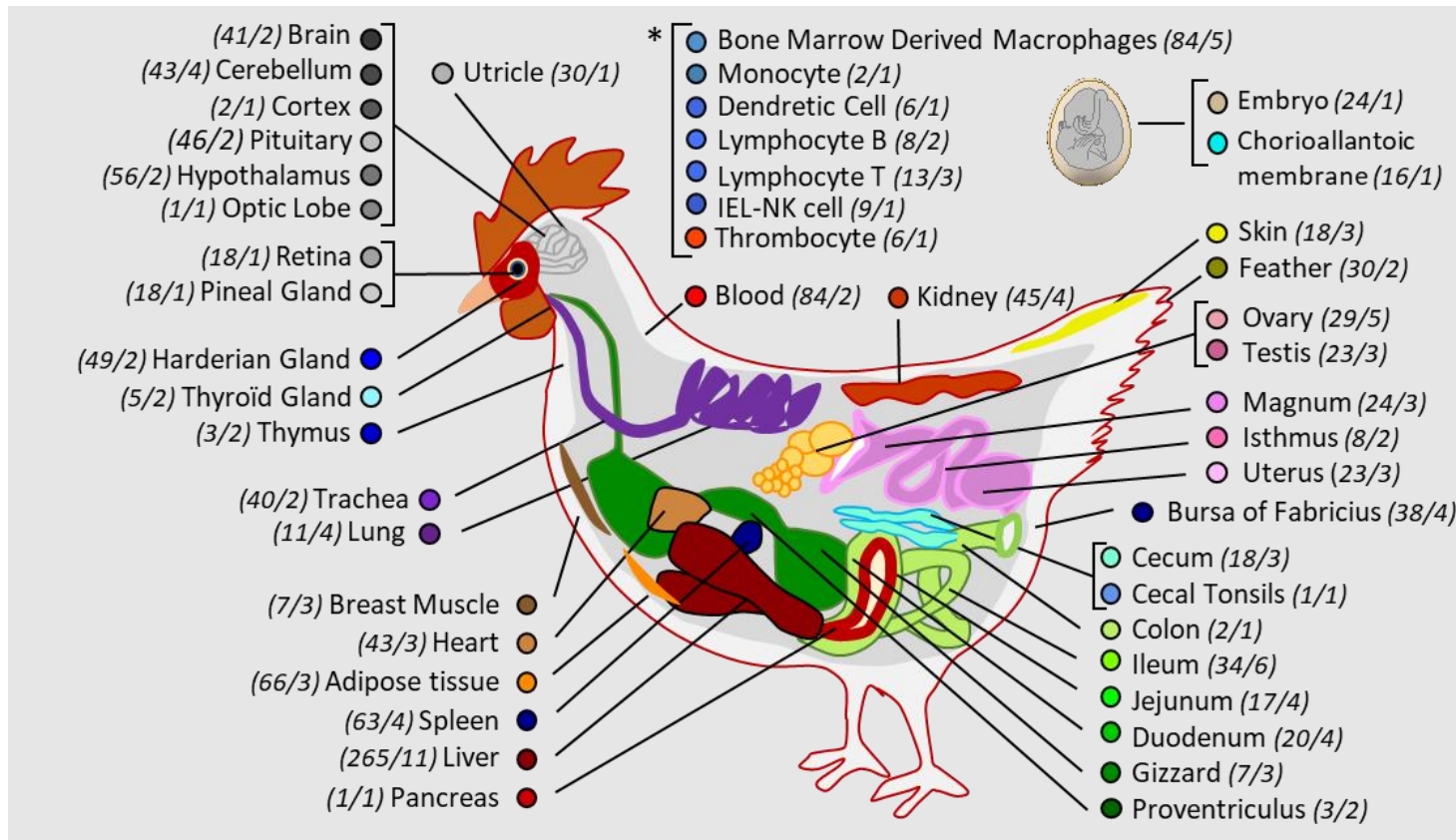
⇒ Using this strategy, we increased the number of **Protein-Coding Genes (PCG)** to 24,000  
& the number of **lncRNA** to 44,000



For each gene locus, we can get the gene occurrence in the 6 databases and their ID equivalence between NCBI

# We tested the reliability of these gene models by quantifying their expression across 47 tissues

=> We analyzed 1400 RNAseq samples from 36 datasets representing the main physiological systems for chickens



In order to easily work on this Gene-Enriched Atlas  
we created **GEGA**: A user-friendly tool

<https://gega.sigena.org/>

**GEGA** *Gallus Enriched Gene Annotation*





# In order to easily work on this Gene-Enriched Atlas we created **GEGA**: A user-friendly tool

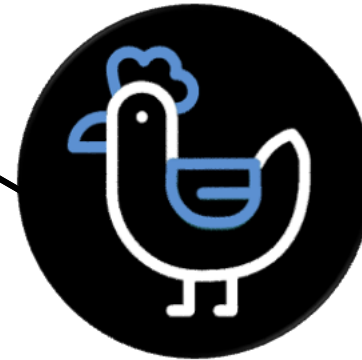
<https://gega.sigenae.org/>

If you are interested in a QTL region

⇒ You can define this region ...

⇒ & display all the genes ...

**GEGA** Gallus Enriched Gene Annotation



⇒ You can then get various genomics & functional information related to these genes

Ex.: Body fat reserve QTL on Chr.2

Define a region

Chr.:2

E.g. Focus on a region or a QTL

Region of interest

90.5Mb 91026526 AX-76180548 91026527 91.5Mb

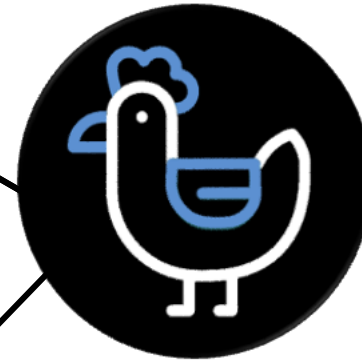
gnid	gnName	gnName_full	chr	start	end	strai	source	version	gnBiotype
LOC428513	TSHZ1	teashirt zinc	2	90,958009	91,012602	-	Ncbi	v106	protein_cod
LOC421010	ZADH2	zinc binding	2	91,014325	91,020611	+	Ncbi	v106	protein_cod
LOC421011	ZNF407	zinc finger	2	91,020611	91,020611	+	Ncbi	v106	protein_cod
DAVISGALGC	DAVISGALGC						s	v1	protein_cod
...	...							...	...
LOC1125318	LOC1125318							v106	protein_cod
DAVISGALGC	DAVISGALGC						s	v1	protein_cod
LOC421012	CNDP1	carboxypeptidase	2	91,437776	91,455226	-	Ncbi	v106	protein_cod
INRAGALG00	INRAGALG00	NA	2	91,458024	91,458225	-	Inra	v2018	protein_cod
LOC421013	CNDP2	CNDP dipept	2	91,470578	91,484819	-	Ncbi	v106	protein_cod
LOC421014	FAM69C	family with s	2	91,498762	91,508921	+	Ncbi	v106	protein_cod
LOC421015	C18orf63	chromosome	2	91,532276	91,554368	-	Ncbi	v106	protein_cod
LOC414798	CYB5A	cytochrome b	2	91,554619	91,565928	+	Ncbi	v106	protein_cod
LOC421016	TIMM21	translocase c	2	91,604515	91,611151	-	Ncbi	v106	protein_cod
LOC421017	FBXO15	F-box protei	2	91,609665	91,633896	+	Ncbi	v106	protein_cod

17 "PCG" genes  
+ 65 "LNC" genes

# In order to easily work on this Gene-Enriched Atlas we created **GEGA**: A user-friendly tool

<https://gega.sigenae.org/>

## **GEGA** Gallus Enriched Gene Annotation



If you are interested in a QTL region

⇒ You can define this region

⇒ & display all the genes

**Functional Ontologies**  
(GO terms / phenotype terms)



**Search**

Filter on specific field(s)

⇒ *Select the genes with ontologies that are consistent with the trait of interest*

**Ex.: Body fat reserve QTL on Chr.2**

**Define a region**

Chr.:2

E.g. Focus on a region or a QTL

Region of interest

90.5Mb 91026526 AX-76180548 91026527 91.5Mb

gnid	gnName	gnName_full	chr	start	end	strai	source	version	gnBiotype
LOC428513	TSHZ1	teashirt zinc	2	90,958009	91,012602	-	Ncbi	v106	protein_codi
<b>LOC421010</b>	<b>ZADH2</b>	zinc binding	2	91,014325	91,020541	-	Ncbi	<b>v106</b>	<b>protein_codi</b>
LOC421011	ZNF407	zinc finger	2	91,020541	91,020541	-	Ncbi	v106	protein_codi
DAVISGALGC	DAVISGALGC						s	v1	protein_codi
...	...							...	...
LOC1125318	LOC1125318							v106	protein_codi
DAVISGALGC	DAVISGALGC								protein_codi
LOC421010									n_codi
INRAGA									n_codi
LOC421010									n_codi
LOC421010									n_codi
LOC421010									n_codi
LOC421010									n_codi
<b>LOC414798</b>	<b>CYB5A</b>	cytochrome	<b>2</b>	<b>91,554619</b>	<b>91,565928</b>	<b>+</b>	<b>Ncbi</b>	<b>v106</b>	<b>protein_codi</b>
LOC421016	TIMM21	translocase c	2	91,604515	91,61151	-	Ncbi	v106	protein_codi
LOC421017	FBXO15	F-box proteir	2	91,609665	91,633896	+	Ncbi	v106	protein_codi

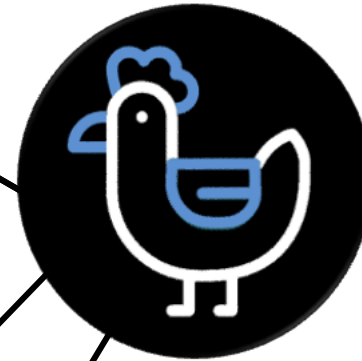
**17 "PCG" genes  
+ 65 "LNC" genes**

⇒ **2 genes CYB5A & ZADH2**  
related to "fat" or "lipid" terms

# In order to easily work on this Gene-Enriched Atlas we created **GEGA**: A user-friendly tool

<https://gega.sigena.org/>

## **GEGA** Gallus Enriched Gene Annotation



If you are interested in a QTL region

- ⇒ You can define this region
- ⇒ & display all the genes

Functional Ontologies  
(GO terms / phenotype terms)

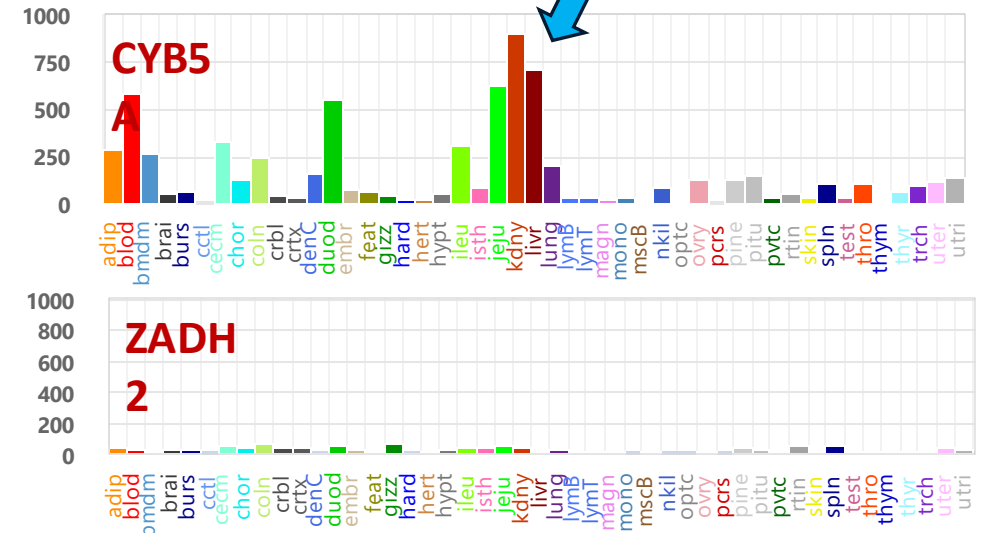
Expressions across 47 tissues

⇒ *Prioritize the genes expressed in a tissue that is consistent with the trait of interest*

With the possibility to display **Graphical Views**

Ex.: Body fat reserve QTL on Chr.2

The gene **CYB5A** highly expressed in the liver (key organ of lipid synthesis)

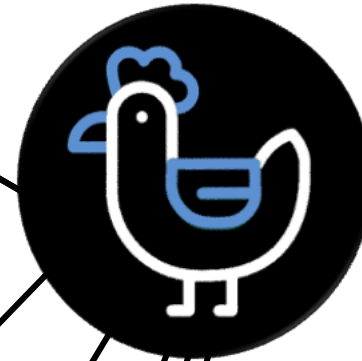


the 2 genes **CYB5A** & **ZADH2** previously selected



# In conclusion...

## GEGA *Gallus Enriched Gene Annotation*



If you are interested in a QTL region

⇒ You can define this region ...

⇒ & display all the genes ...

**Functional Ontologies**  
(GO terms / phenotype terms)

**Expressions** across 47 tissues

**Genomics orientation** (convergent / divergent / antisense) btw 2 neighboring genes

**Orthologous genes** (mouse & human)

**Co-expressions** in chicken & human

This tool should be useful  
for the GeroNIMO project  
to explore the multiple QTL regions  
In chicken



Philippe Bardou  
INRAE  
(SIGENAE)



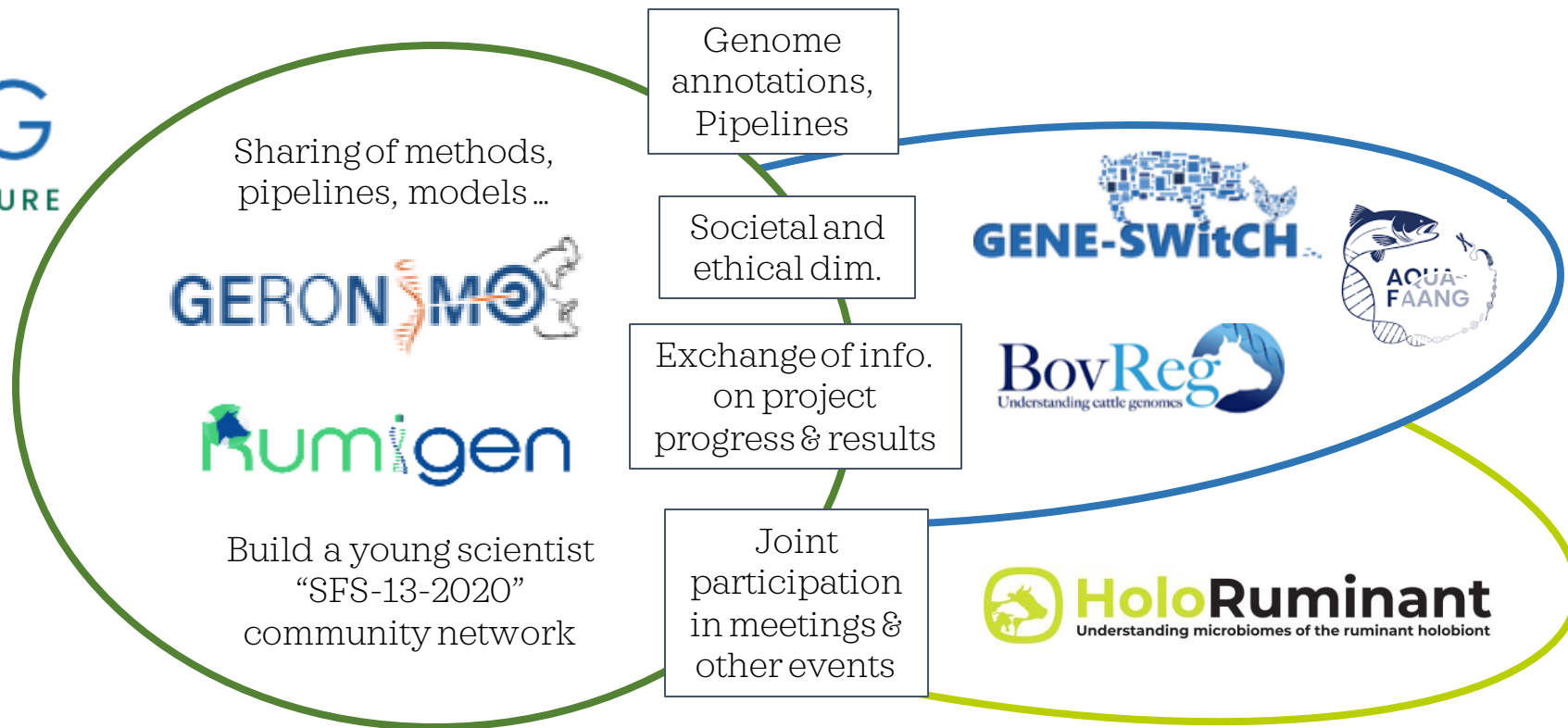
Fabien Degalez  
PhD student  
UMR Pegase  
INRAE-Institut Agro

For more details about this atlas  
=> see the poster of F. Degalez

# Thank you for your attention



New knowledge on functional genomics



**GEGA** Gallus Enriched Gene Annotation <https://gega.siginae.org/>

