AQUA-FAANG: Empowering Aquaculture Research and Innovation By Genome Functional Annotation

Dr Peter Harrison. European Bioinformatics Institute, EMBL-EBI
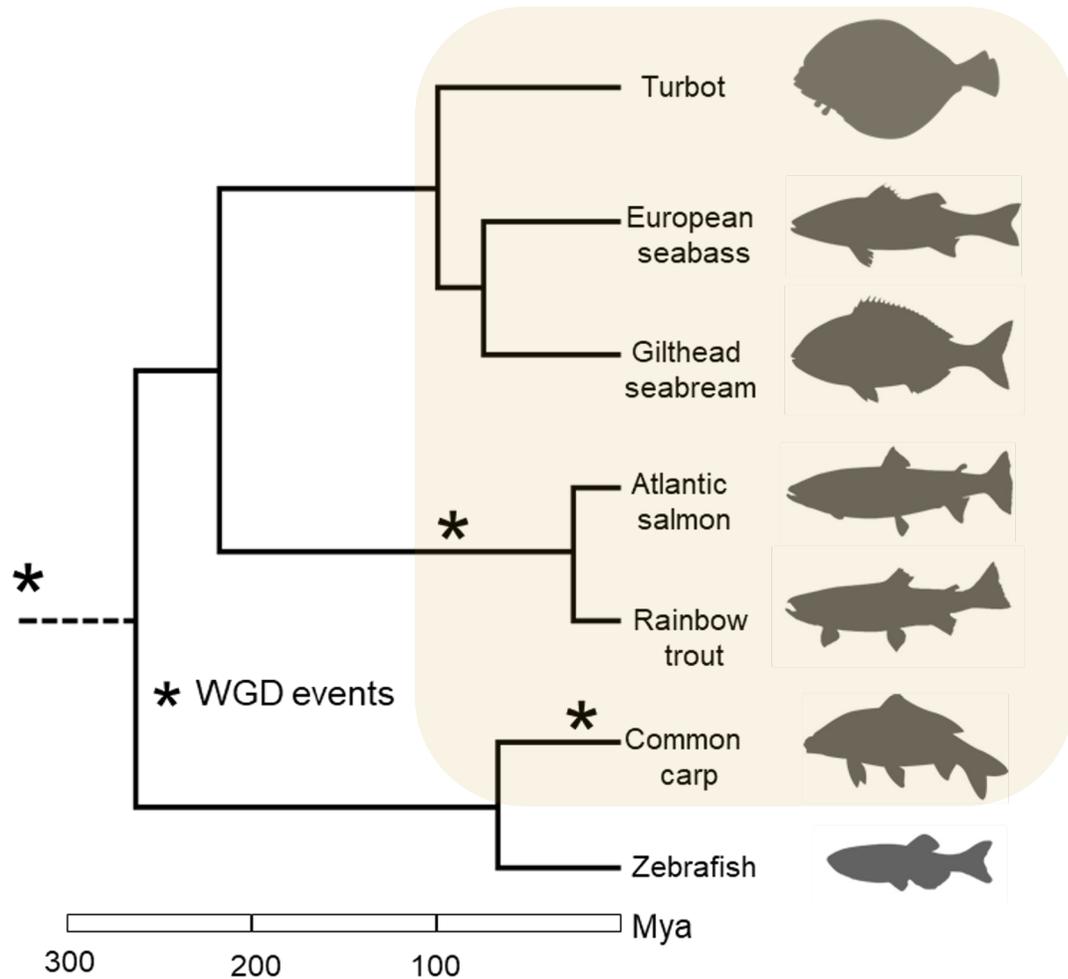
On behalf of AQUA-FAANG consortium

Functional Annotations of Animal Genomes (FAANG) Workshop – PAG31 – San Diego – Jan 16th 2024

AQUA-FAANG

Aim: produce genome-wide functional annotation maps for six commercially important fish species

- Funded at 6 million Euros by H2020
- May 2019 to Oct 2023
- Coordinated by Sigbjørn Lien & Dan Macqueen
- One of the six EuroFAANG projects

AQUA-FAANG Consortium

14 academic partners & 10 industry partners across 9 countries

Specific objectives

AQUA-FAANG - functional annotation

Regulatory elements, chromatin epigenetic state and gene expression

- Standard samples across species - *different life stages, tissues from both sexes, and responses to immune challenge*

DevMaps ⎯ Six standardized stages of embryogenesis (all assays) plus at least six additional stages (RNA-Seq)

♀♂ **BodyMaps** ⎯ Tissue panel in both sexes, sexually immature and mature:
- Liver, brain, gill, intestine, skeletal muscle, head kidney, ovary and testis

Histone modifications
H3K27ac (Active promoter/ Enhancer)
H3K4me1 (Active enhancer)
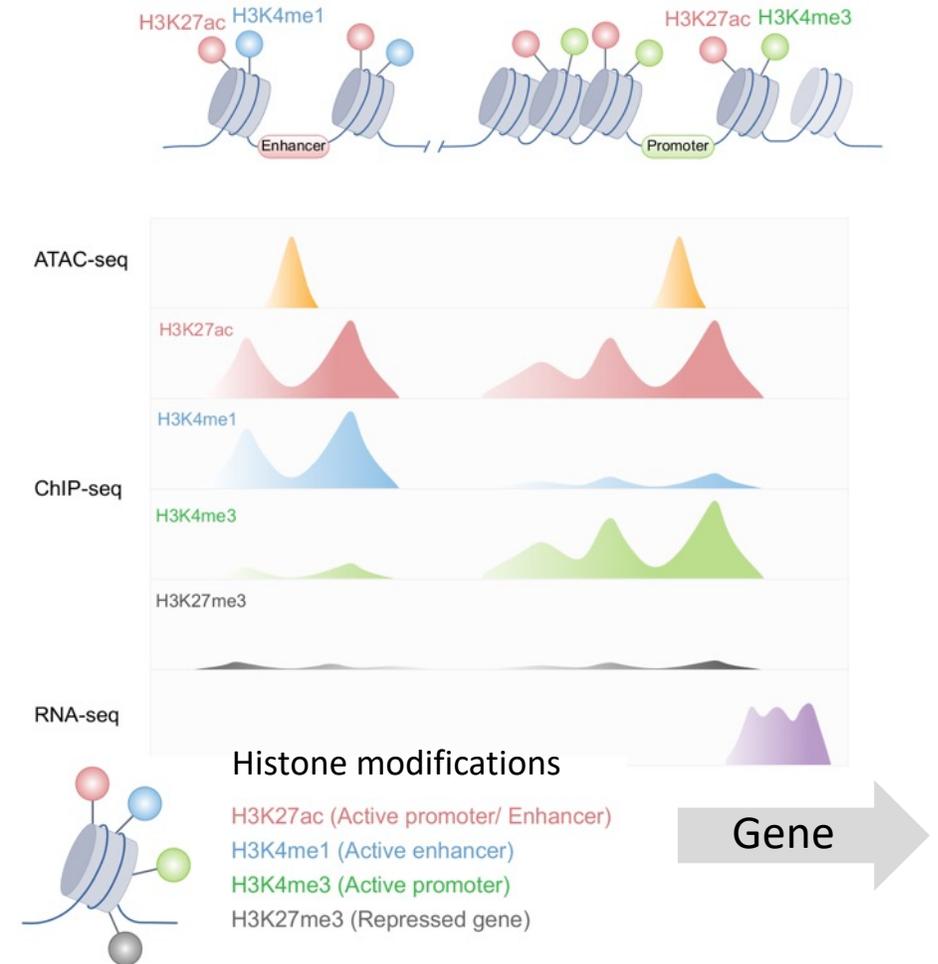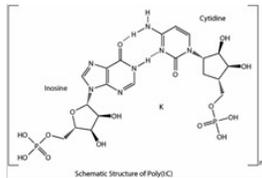H3K4me3 (Active promoter)
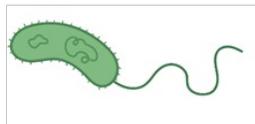H3K27me3 (Repressed gene)

Gene

Image credit: Pooran Dewari

AQUA-FAANG - functional annotation

Regulatory elements, chromatin epigenetic state and gene expression
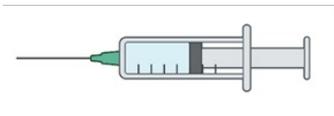
- Standard samples across species - *different life stages, tissues from both sexes, and responses to immune challenge*

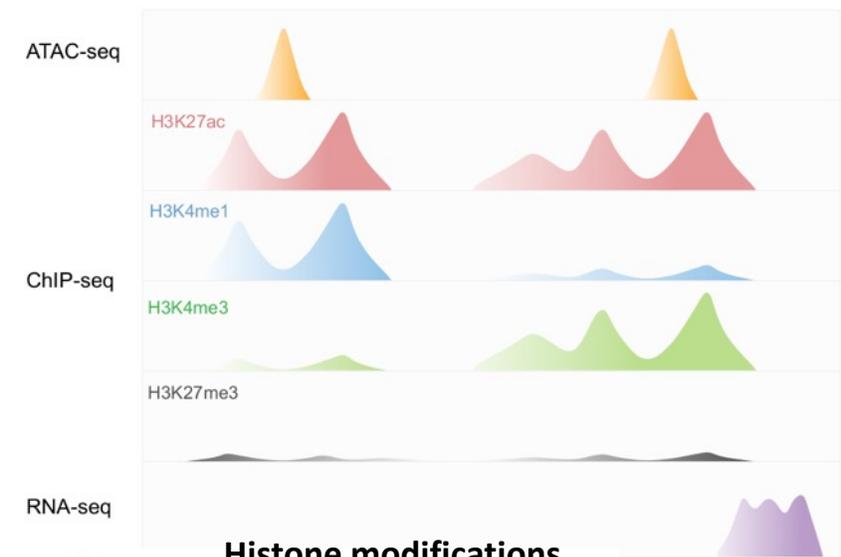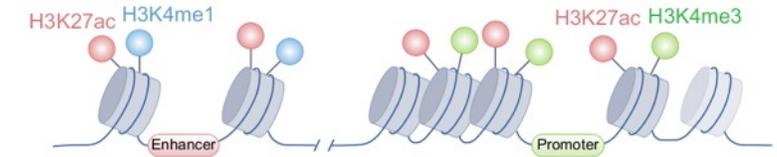**Poly I:C**

Inactivated Vibrio

In vivo

In vitro

ImmunoMaps

- Head kidney samples - primary haemopoietic organ
- Poly I:C (viral mimic)
- Inactivated *Vibrio* (pro-inflammatory response)
- PBS controls
- *in vivo* I.P. injection and *in vitro* leucocyte cultures

ATAC-seq

ChIP-seq

RNA-seq

Histone modifications

H3K27ac (Active promoter/ Enhancer)
H3K4me1 (Active enhancer)
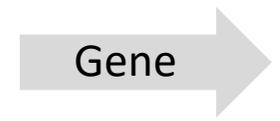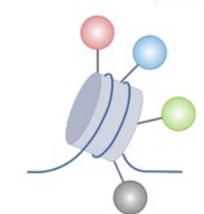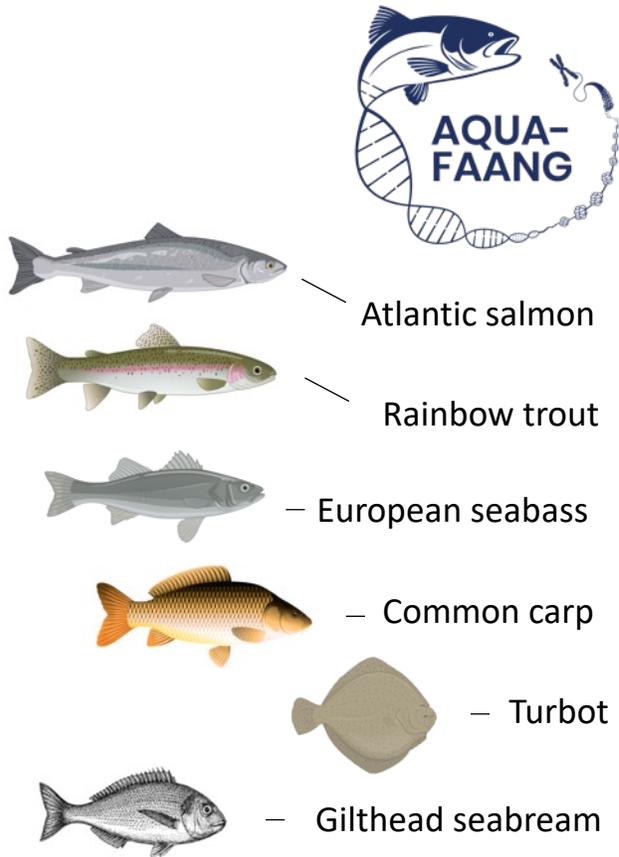H3K4me3 (Active promoter)
H3K27me3 (Repressed gene)

Gene

Image credit: Pooran Dewari

Data produced across species

	BodyMaps			DevMaps			ImmunoMaps		
	RNA-Seq *mRNA + small*	ATAC-Seq	ChIP-Seq	RNA-Seq *mRNA + small*	ATAC-Seq	ChIP-Seq	RNA-Seq *mRNA + small*	ATAC-Seq	ChIP-Seq
Atlantic salmon	108	60	299	71	21	106	54	36	78
Rainbow trout	114	60	239	71	15	95	122	68	52
European seabass	116	60	257	75	18	61	54	36	126
Common carp	108	60	240	67	18	68	59	36	123
Turbot	108	60	194	62	18	56	54	36	114
Gilthead seabream	116	60	166	62	18	45	88	36	119

Approx. 4,700 sequencing datasets produced across species as basis for functional annotations

- > 50 trillion DNA base pairs of new information on genome function
- Many partners and individuals involved – **too many to acknowledge individually, amazing team effort**

Data analysis & sharing

Data used to **update gene model predictions** (RNA-Seq) and produce **novel annotations of regulatory features** (ATAC-Seq & ChIP-Seq) in **Ensembl genome browser**

Peter Harrison   Garth Ilsley   Fergal Martin

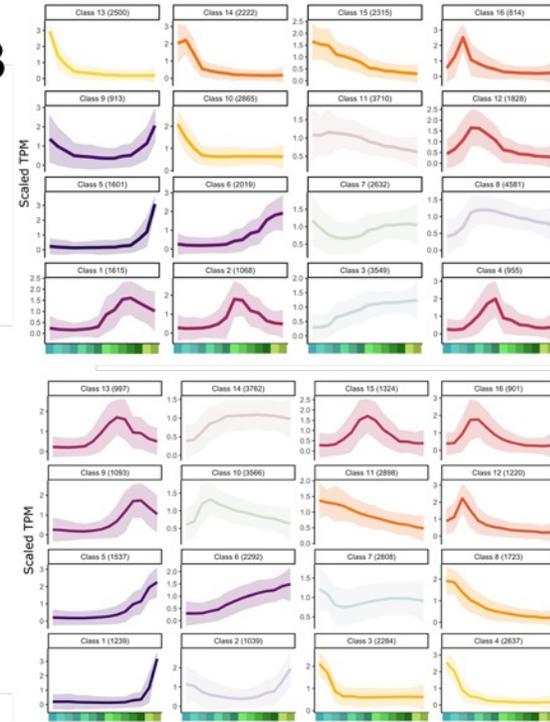**Data and metadata available on FAANG data portal**

- Partners performed own data analysis and annotations using nf-core pipelines - *for standardization and to harmonise with other EuroFAANG projects*

- Large up-step in competence in bioinformatics for regulatory data

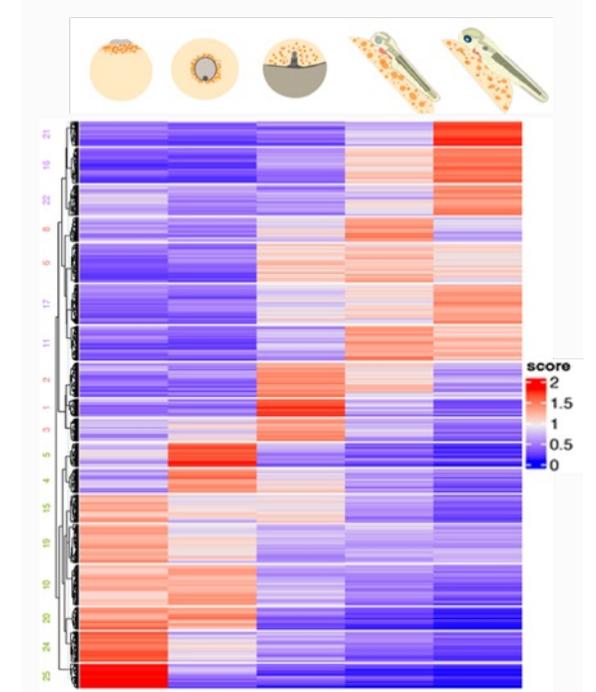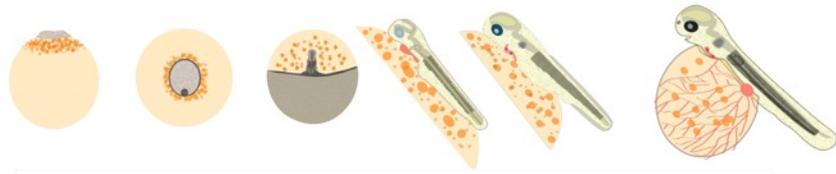# Highlights: revealing functional activity in fish genomes
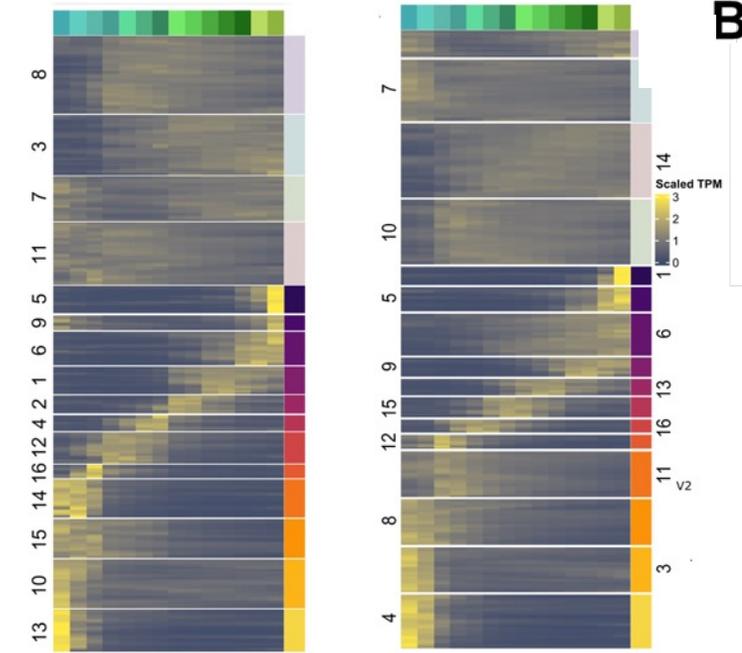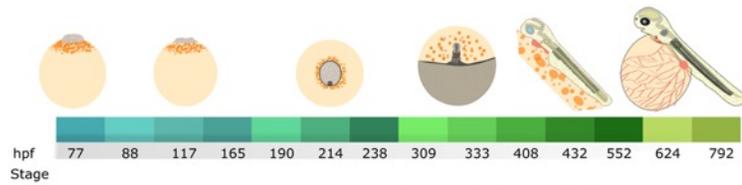
Transcriptome

Chromatin accessibility

Self organizing map clustering: expression & chromatin accessibility

Highlights: revealing functional activity in fish genomes

Embryogenesis stages   Image credit: Diego Perojil

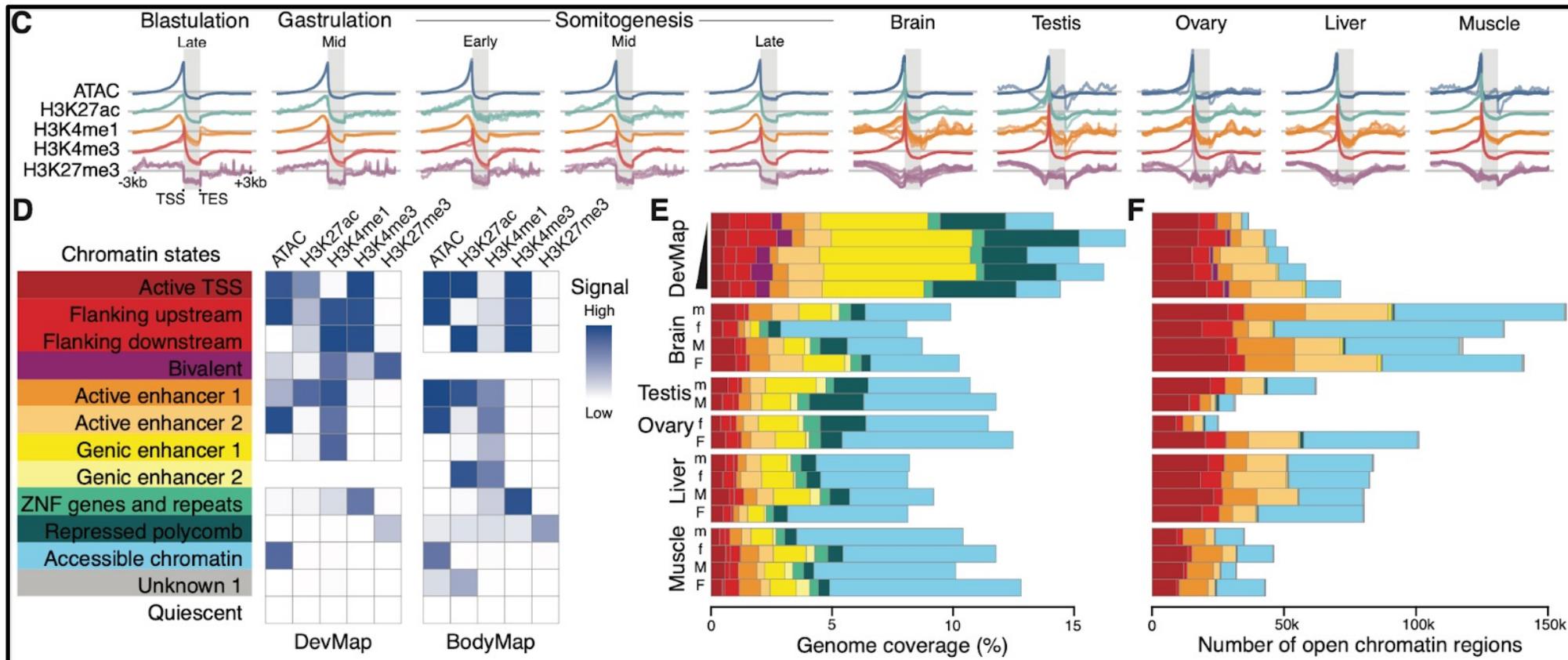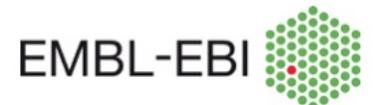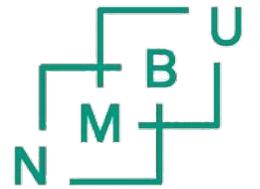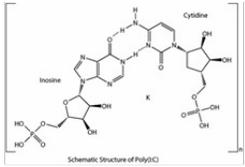**Tissue panel from immature & mature fish**

Image credit: Gareth Gillard

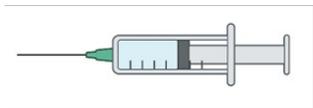# Highlights: revealing functional activity in fish genomes

Poly I:C

Inactivated *Vibrio*

In vivo

In vitro

Regulatory regions were activated in response to pathogen signals -> variants overlapping such elements are likely to influence disease resistance phenotypes
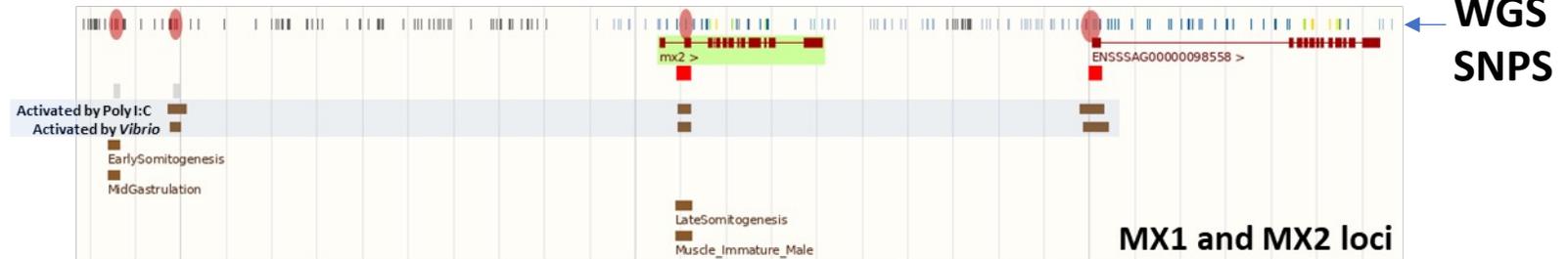
- Strong activation of regulatory elements surrounding interferon pathway genes in response to poly I:C (viral mimic)

Mx Proteins: Antiviral Gatekeepers That Restrain the Uninvited
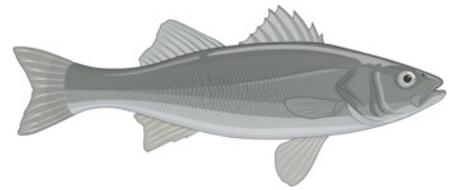
Judith Verhelst, Paco Hulpiau, Xavier Saelens

Inhibition of Infectious Pancreatic Necrosis Virus Replication by Atlantic Salmon Mx1 Protein

Rannveig Larsen, Torunn P. Røkenes, and Børre Robertsen*

WGS SNPS

mx2 >

ENSSSAG00000098558 >

Activated by Poly I:C
Activated by *Vibrio*

EarlySomitogenesis

MidGastrulation

LateSomitogenesis

Muscle_Immature_Male

MX1 and MX2 loci

Ensembl regulatory build (ATAC-only)

■ promoter   ■ open chromatin   ■ active enhancers (UEDIN/NMBU/UNIABDN ChromHMM)

● SNPs overlapping active enhancers

Analysis/image credit: Shahmir Naseer and Sam Martin

1495
UNIVERSITY OF ABERDEEN

THE UNIVERSITY OF EDINBURGH

Highlights: functional annotation to understand the genetic basis of disease resistance
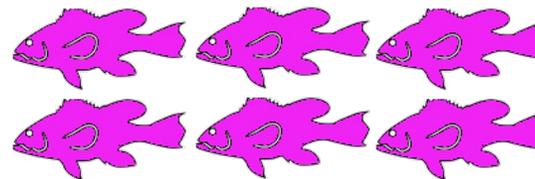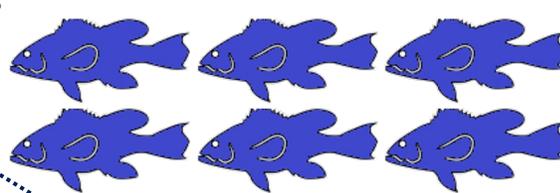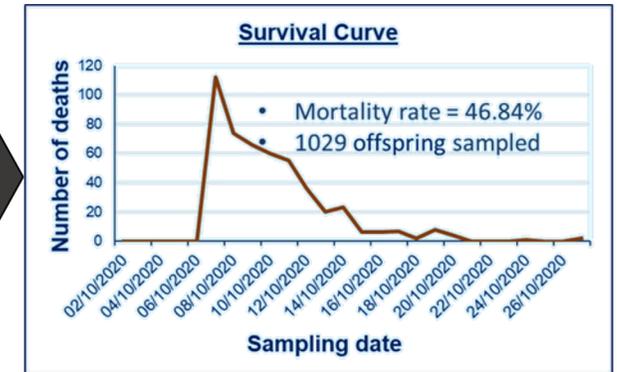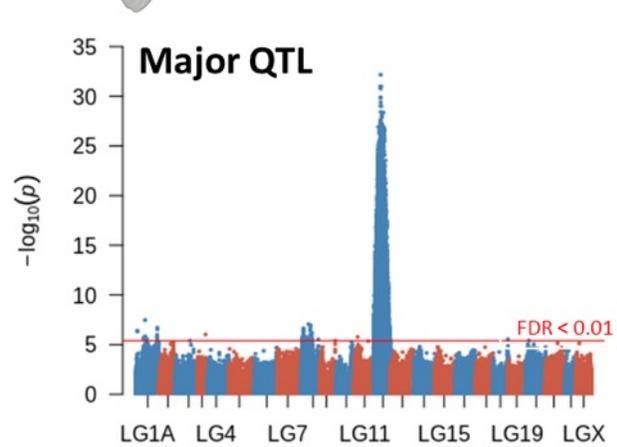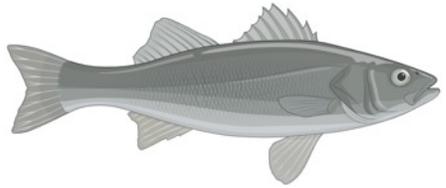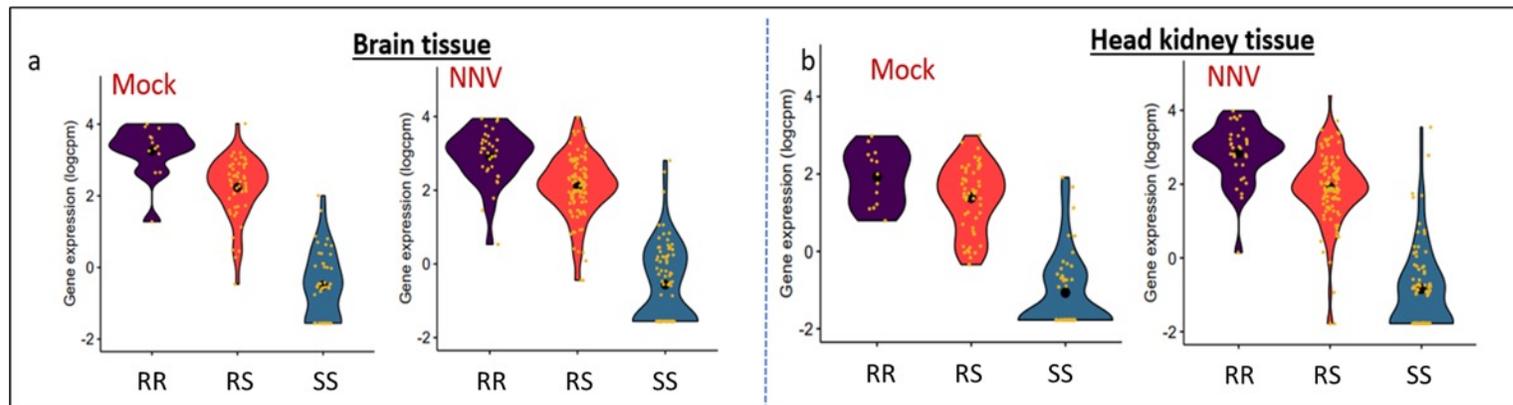

Exemplar study to uncover the functional genetic architecture of viral nervous necrosis resistance

Highlights: functional annotation to understand the genetic basis of disease resistance
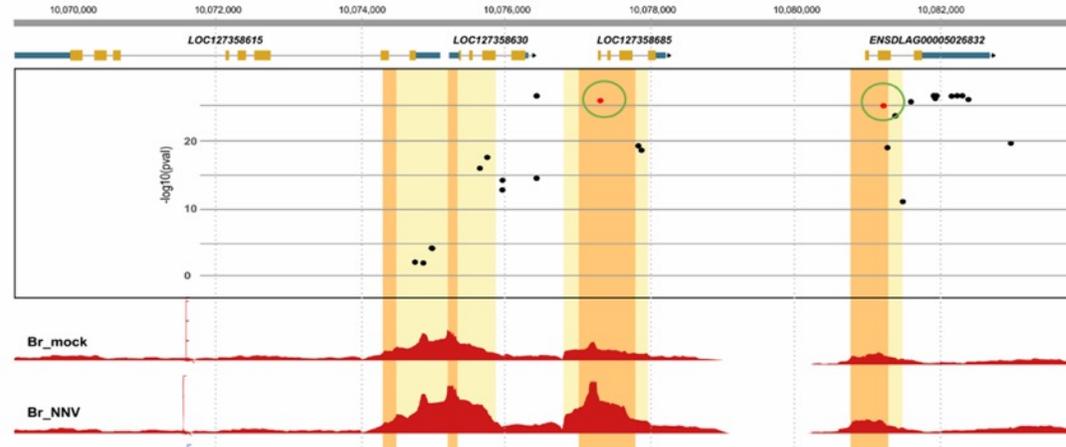
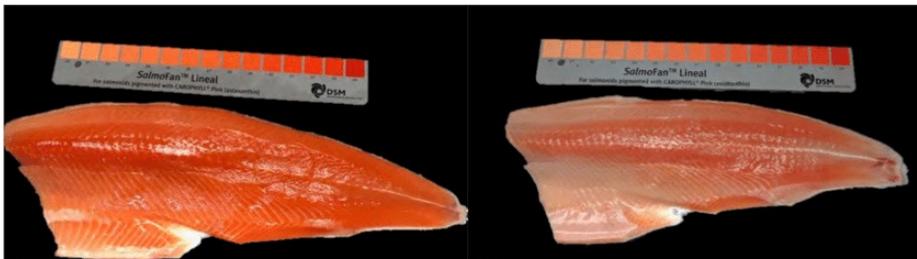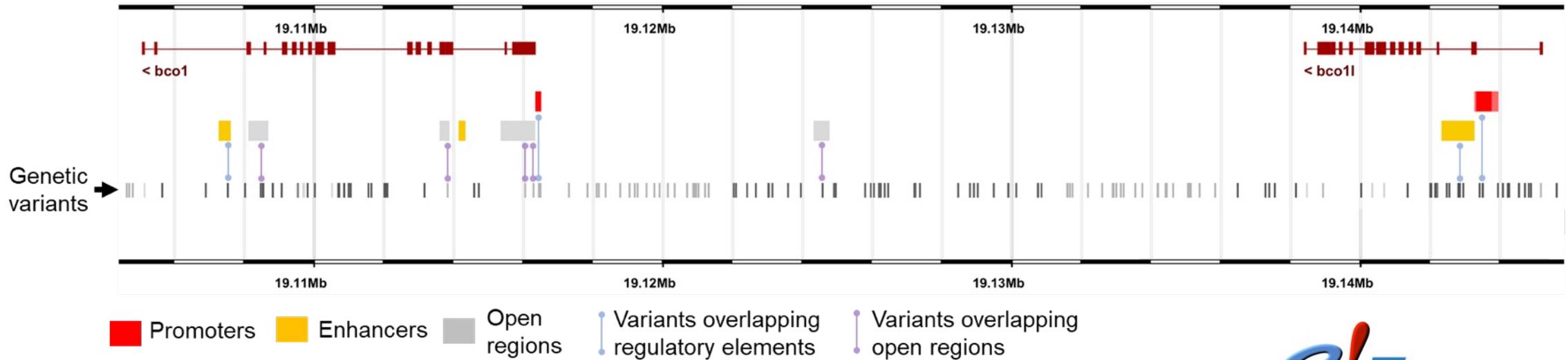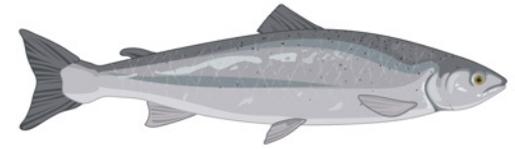*Two SNPs in open chromatin of two copies of ifi27l2a*

eQTL analysis - ifi27l2a RNA levels associated with resistance in virus target tissues

Exemplar study to uncover the functional genetic architecture of viral nervous necrosis resistance

Highlights: prioritizing variants

e.g. 40 kb region impacting flesh colour

Figure adapted from Johnston et al. 2024. AQUA-FAANG whitepaper - In Revision (Aquaculture)

Exploiting functional annotation for more precise breeding

Deeper annotation will increase resolution to identify causative variants

Figure adapted from Johnston et al. 2024. AQUA-FAANG whitepaper - **In Revision (Aquaculture)**
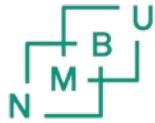
Summary of achievements

1. Developed standardized functional annotation protocols for six major aquaculture species and shared the resulting data and expertise

2. Generated and interpreted functional annotation maps for both healthy and immune activated states to identify functional regions across tissues, life stages, sexes, and after viral or bacterial challenge

3. Exploited functional annotation to help predict functionality of variants responsible for resistance to aquaculture pathogens and used project results to improve breeding approaches in aquaculture

4. Developed resources and methods for comparative analysis of functional annotation datasets, to reveal genomic elements conserved across species and in relation to WGD events in fish evolution

5. Actively engaged many project stakeholders in aquaculture sector through a comprehensive program of communication, dissemination, and exploitation activities

Thank you for your attention!