BOOK OF ABSTRACTS
AQUA-FAANG
Final Conference

Edinburgh

11th-13th October 2023

The AQUA-FAANG project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923.
www.aqua-faang.eu

Meeting Agenda

Day 1: Wednesday 11th October

9:00-9:10 Welcome and introduction: A short history of AQUA-FAANG
Sigbjørn Lien, NMBU

Session 1: Improving the functional annotation of farmed fish genomes, 9:10-12:20

9:10-9:30 Development of functional annotation assays
Matthew Peter Kent, NMBU

9:30-9:50 Ensembl regulation's annotation pipeline
Gabriela Merino, EMBL-EBI

9:50-10:10 Using Ensembl's regulatory annotation
Garth Ilsley, EMBL-EBI

Coffee break 10:10-10:40

10:40-11:00 Functional miRNA annotation in teleost fish
Julien Bobe, INRAe

11:00-11:10 Preliminary analyses of lncRNAs and muscle related ohnologues between AQUAFAANG species
Daniel Garcia de la serrana, University of Barcelona

11:10-11:30 An Atlas of regulatory elements and structural variants in Turbot: potential implications for improving farming
Paulino Martinez, USC

11:30-12:20 Celebrating 10 years of FAANG - From FAANG to Fork: Highly annotated genomes as resources to improve farmed animal production
*Guest speaker: Emily Clark, University of Edinburgh - **the EuroFAANG Research Infrastructure Project***

Lunch 12:20-13:20

Session 2: The dynamic functional regulation of farmed fish genomes 13:20-17:00

13:20-13:40 Evolution of duplicated genome regulation in salmonids
Marie-Odile Baudement, NMBU

13:40-14:00 Gene expression and regulation across salmonid ontogeny
Diego Perojil Morata, Uni Edinburgh

14:00-14:20 Comparative regulomics gives insights into the conservation and evolution of regulatory elements following whole genome duplication in salmonids
Manu Kumar Gundappa, Uni Edinburgh

Meeting Agenda

14:20-14:40 Linking divergence of salmonid gene expression to regulation
Gareth Gillard, NMBU

14:40-15:00 Differences in transcription initiation in early embryogenesis between Cyprinid species
Damir Baranasic, Imperial College London

Coffee break 15:00-15:30

15:30-15:50 Exploring the mechanisms behind allotetraploid genome regulation using carp as a model
Ada Jimenez-Gonzalez, Uni Birmingham

15:50-16:20 Genomic and cellular insights into antiviral responses and viral disease resistance in salmonid fishes
Thomas Clark, INRAe/Uni Aberdeen

16:20-16:40 Comparative regulomics in flatfish: from turbot to the main farmed Pleuronectiformes
Juan Rubiolo, USC

16:40-17:00 Dynamic gene expression and regulation during gilthead sea bream development
Elena Sarropoulou, HCMR

Drinks reception 18:00-19:00
Playfair Library, Old College, South Bridge, Edinburgh

Dinner 19:00
Playfair Library, Old College, South Bridge, Edinburgh

Day 2: Thursday 12th October

Session 3: Functional genomic basis for immune responses and disease resistance in farmed fish, 9:00-13:00

9:00-9:20 Genome functional annotation of host defense response in gilthead sea bream (*Sparus aurata*) through chromatin accessibility and differential gene expression assays
Costas Tsigenopoulos, HCMR

9:20-9:40 Transcriptome and chromatin landscape of European seabass immune response to viral-like stimulation
Serena Ferraresso, UNIPD

9:40-10:00 Functional genomic architecture of viral nervous necrosis disease resistance in farmed European seabass
Robert Mukiibi, University of Edinburgh

Meeting Agenda

10:00-10:45 Decoding enhancer function: from the nucleosome to the nucleus
Guest Speaker: Wendy Bickmore, University of Edinburgh

Coffee break 10:45-11:20

11:20-11:40 Transcriptional differences in CyHV-3 response between resistant and susceptible common carp (Cyprinus carpio) crossings
Lukasz Napora-Rutkowski, ZIGR

11:40-12:00 Symmetric expression of ohnologs encoding conserved antiviral responses in tetraploid common carp suggest absence of subgenome dominance after whole genome duplication
Hendrik-Jan Megens, Wageningen University

12:00-12:20 Multiomics reveals the genomic regulatory landscape underlying the antiviral response in Atlantic salmon
Shahmir Naseer, University of Aberdeen

12:20-12:40 In vitro mutant models for functional characterization of genes of the type I IFN pathway in salmonids
Pierre Boudinot, INRAe

12:40-13:00: A turbot immune map: comparative response to bacteria and virus stimulation regarding previous genomics data on industrial pathogens
Oscar Aramburu, USC

Lunch 13:00-14:00

14:00-15:15 AQUA-FAANG 2.0: what's next? Group discussion
Please note: this will involve simultaneous discussions among 5-6 breakout groups. We will therefore not stream this session for the online audience. The stream for day 2 ends at lunch time, 13:00 CET.

Pub quiz 16:30-18:30
Held at the Salisbury Arms, approximately 5 min walk from the conference centre

Free Evening - Self-arranged dinner and time to explore Edinburgh

Meeting Agenda

Industry Day: Friday 13th October

Session 4: AQUA-FAANG relevance to industry, 9:00-11:10AM
9:00-9:15 Basic overview of AQUA-FAANG and its potential applications in industry
Dan Macqueen, University of Edinburgh

9:15-9:30 Ensembl gene annotation, regulation and variant effect prediction for aquaculture
Peter Harrison, EMBL-EBI

9:30-10:10 Accounting for overlapping annotations as biological priors in genomic prediction models of complex traits
*Andrea Rau, INRAe - **The GENE-SWitCH Project***

10:10-10:50 Functional genomics and selective breeding in aquaculture: implications from the AQUA-FAANG project
Speakers: Diego Robledo, Sara Faggion, Robert Mukiibi

10:50-11:10 Developing a flexible, low cost, multifunctional genotyping solution for selective breeding in aquaculture
Rachael Wilbourn, Xelect

Coffee Break 11:10-11:40

Session 5: Perspectives from industry, 11:40-13:00

11:40-12:00 A way out of black box genomic selection
*Antti Kause, LUKE - **The AquaIMPACT project***

12:00-12:20 Advancing selective breeding in aquaculture through the functional annotation of fish genomes
Mark Looseley, Xelect Ltd.

12:20-12:40 Regulatory Elements in Genomic Selection
Tim Knutsen, AquaGen

12:40-13:00 Potential use of AQUA-FAANG results to develop different breeding programs in diverse aquaculture companies, increasing their efficiency, profitability and sustainability.
Adrian Millan, GeneAqua

Lunch 13:00-14:00

14:00-15:00 Panel Discussion.
Covering the main talking points of day 3 and future perspectives for aquaculture sector
Chair: Ashie Norris, Mowi

15:00-15:10 Final words and acknowledgements

Abstracts

Day One, Wednesday 11th October

Development of functional annotation assays

Mattew Peter Kent, NMBU

The AQUA-FAANG project is composed of various ambitious Work-Packages (WPs) and tasks with a high degree of interdependence that collectively seek to improve our understanding of the functional genomic landscape of six important fish species. WP1 has been the starting point for our journey of discovery and was responsible for generating the vast amount of raw data that has been analysed throughout the project. Specific goals have included (i) identifying appropriate lab protocols suitable for generating functional annotation data and adapting these to work in diverse fish species and tissue types, (ii) training and enabling partners in the project to be able to perform lab work thereby developing skills and experience, (iii) generating thousands of novel data sets that collectively describe the functional genomic landscape at different life stages, in different tissues and under different immunologically stressed conditions, in all six focus species, and (iv) establish guidelines for both standardized sampling and sample detail recording so that data is accompanied with information describing its source. Achieving all of these goals has been one of the most resource demanding aspects of the project involving dozens of people and a significant portion of the budget, but it has been successful and resulted in a comprehensive and unparalleled collection of unique data that informs our understanding of gene-regulation in the target species and, by comparative extension, other fish species. Protocols and data are publicly accessible, and lay the foundation for knowledge based advancement of aquaculture and general fundamental research into teleosts.

Abstracts

Ensembl's regulatory annotation pipeline

Gabriela Alejandra Merino[1], Paulo R Branco Lins[1], Malcolm Perry[1], David Urbina-Gómez[1], Garth Ilsley[1] and Peter W. Harrison[1]

1. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, CB10 1SD, Cambridge, UK

Ensembl (https://www.ensembl.org/) is a widely used genome browser that provides tools for assisting the scientific community in genome interpretation. Ensembl's regulatory annotation identifies features that might regulate and control gene expression. The AQUA-FAANG consortium has produced high-quality data from multiple functional assays for some of the economically most important fish species within European aquaculture. Here we present the new Ensembl primary analysis pipeline and regulatory build we developed for AQUA-FAANG.

Our primary analysis pipeline processes ATAC-seq and ChIP-seq data from alignment to peak calling. We developed a strategy to mask repetitive genomic regions using control ChIP-seq data. For peak calling, we use Genrich which provides parameters that we optimised for each type of assay. We also developed a postprocessing step to detect low-quality experiments by analysing a set of summary statistics including the number of peaks and the percentage of exclusive peaks per experiment.

Our new regulatory build is based on open chromatin peaks detected across the wealth of data generated by the AQUA-FAANG project. It also uses annotated transcription start sites, exons, and histone marks to annotate promoters, enhancers, and open chromatin regions. Initial annotation for Atlantic salmon, Turbot, and European seabass was released in Ensembl 110 (July 2023). Improvements on these annotations and new annotations for Rainbow trout and Common carp will be published in Ensembl 111 (due to be released in November 2023).

Abstracts

Using Ensembl's regulatory annotation

Garth Ilsley[1], Gabriela Alejandra Merino[1], Paulo R Branco Lins[1], Malcolm Perry[1], David Urbina-Gómez[1] and Peter W. Harrison[1]

1. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, CB10 1SD, Cambridge, UK

This year Ensembl has added regulatory annotation for AQUA-FAANG fish species in Ensembl 110, with further updates upcoming in Ensembl 111. The annotation identifies candidate regulatory regions in the genome and categorises them as promoters, enhancers or open chromatin regions. The preceding talk from Gabriela Merino titled “Ensembl’s regulatory annotation pipeline” focused on how Ensembl’s annotation is produced; this complementary talk will expand on this overview to focus on its use. It will follow a practical approach, showing how to access and view the data, and how to narrow the search for candidate regulatory regions based on activity across epigenomes. Some guidelines and limitations of interpretation will also be highlighted. These are the first Ensembl regulatory builds for aquaculture, and is a key enhancement to the offering from Ensembl for researchers and industry specialists.

Ensembl is primarily funded by the Wellcome Trust (WT222155/Z/20/Z). The AQUA-FAANG project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement no 817923.

Abstracts

Functional miRNA annotation in teleost fish

Cervin Guyomar[1], Thomas Desvignes[2], Julien Bobe[3]

1. Sigenae, GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326, Castanet Tolosan, France
2. Institute of Neuroscience, University of Oregon, Eugene OR 97403, USA
3. INRAE, LPGP, 35000, Rennes, France

MicroRNAs (miRNAs) are short non-coding RNAs of approximately 22 nucleotides involved in the post-transcriptional regulation of gene expression. While miRNAs are receiving growing attention in various disciplinary fields, researchers face methodological difficulties and, in many cases, lack a comprehensive characterization of miRNA gene annotations. Even though miRNAs are evolutionarily well conserved, miRNA predictions based on mature forms alone often result in an overprediction of putative miRNA genes by various computational tools. In teleost fishes, additional whole genome duplication events further complicate comparisons to mammalian miRNAomes. We recently established the FishmiRNA database(1) (www.fishmirna.org) in which miRNAs of 10 teleost and 2 holostean species could be manually annotated based on RNA-seq reads aligned to the species' genome, the relative positions of 5p and 3p mature miRNA forms, and existing miRNA annotations in closely related species(2). To take this approach to the next level and accommodate the increasing demand for miRNA annotation in various fish species, including within the AQUA-FAANG project, we developed an automatic miRNA gene annotation pipeline using miRNA gene synteny to guide and improve miRNA gene prediction and annotation. This pipeline was first evaluated using species available in the FishmiRNA database, then used to annotate miRNAs from the rainbow trout and seabream AQUA-FAANG body map RNA-seq dataset. In seabream, we were able to comprehensively annotate 298 miRNA genes producing 459 unique mature miRNAs. This number is highly consistent with previously characterized miRNA gene repertoires in teleosts ranging from 275 to 324 genes. In rainbow trout, we annotated 666 miRNA genes producing 666 unique mature miRNA, indicating that most rainbow trout miRNA genes were retained in duplicates following the Salmonid-specific genome duplication, as previously hypothesized(3). This rainbow trout miRNAome annotation represents a major quantitative leap when compared to the previously characterized 384 mature miRNAs produced from at least 280 genes in this species(4).

Abstracts

Preliminary analyses of lncRNAs and muscle related onhologues between AQUAFAANG species

Daniel Garcia de la serrana, University of Barcelona

Long non-coding RNAs are a group of RNAs over 200nt that do not codify for proteins and that can modulate the transcription and translation of genes through multiple mechanisms. However, while orthologues for protein-coding genes can be studied with available tools, the same analyses for lncRNAs are not straight forward. We have used the data generated during the AQUAFAANG project to identify possible conserved lncRNAs between the six species of the project. To that end we used three different approaches: correlation and similarity analyses and predicted interactions. Using correlation analysis with embryonic data we have identified 60 groups of highly correlated lncRNAs (over 0.90 Pearson index) while similarity analysis found over 100 candidates. So far, syntenic analysis between species have failed to identify common genomic surroundings, what it is at odds with an orthologue relationship. We have also tried to identify duplicated lncRNAs that might have been originated during the teleost whole genome duplication. We have identified several duplicated lncRNAs within the teleost genomes that are very likely originated by duplication. So far, the bodymaps data has not been incorporated in the analysis but it will help to better analyse the lncRNAs, conservation and function.

Abstracts

An Atlas Of Regulatory Elements Across Developmental Stages And Tissues In The Turbot (*Scophthalmus maximus*)

Oscar Aramburu[1], Andrés Blanco[1], Belén Gómez-Pardo[1], Christina Kriaridou[2], Zexin Jiao[2], Pooran Dewari[2], Diego Perojil[2], Diego Robledo[1,2], Dan Macqueen[2], Carmen Bouza[1], Paulino Martínez[1]

1. University of Santiago de Compostela, Spain
2. University of Edinburgh, UK

A comprehensive characterization of regulatory elements of the genome across tissues and developmental stages represents an essential resource to understand the genetic basis of productive traits and to identify genetic variants potentially associated with valuable phenotypes to improve genomic predictions in breeding programs. We identified and characterized regulatory elements in the turbot genome by integrating genome-wide ATAC-seq (open chromatin regions), CHIP-seq (H3K4me3-active promoter regions, H3K27ac-enhancer and promoter regions, H3K27me3-Polycomb repression, H3K4me1-enhancer and promoter regions) and RNA seq datasets from 54 pools of 6 embryonic stages (blastula, gastrula, early-segmentation, mid-segmentation, late-segmentation and prehatching) and 6 immature and 6 mature tissues (gonad, brain, liver, head kidney, gill and muscle) tissues. In total, we identified 10 distinct chromatin states among embryonic stages and tissues and annotated > 100,000 highly reproducible regions. Among the 205,000 open chromatin regions mapped, ~53,000 overlapped promoter regions and ~75,000 intergenic and intronic regions in at least one sample, suggestive of the presence of regulatory elements. Correlation of read counts between all sample pairs in the identified genomic regions suggested the epigenetic mark as the main responsible, while samples seemed not to be so relevant. Enhancers identified where explored in the different samples and the binding motifs retrieved to characterize transcription factors by stage and tissue. Also, we characterized the miRNAome of turbot from: i) pools of samples of 12 embryonic stages, ii) eight organs of juveniles and adults, males and females. A total of 260 miRNAs and their mature forms were identified and mapped in the turbot genome, and their expression characterized across stages and tissues. PCA showed a major differentiation of the brain from the remaining tissues across the 1st component, which were separated from gonads to head kidney across the 2nd component. Furthermore, 15x whole genome resequencing of 54 turbot adults from a representative sample of farm broodstock enabled the detection of thousands of SNPs and structural variants overlapping with regulatory elements, which are being evaluated for their impact on gene expression for a more precise phenotypic evaluation and breeding value estimation. This information is being applied in turbot breeding programs for improving growth and resilience to scuticociliatosis, a systemic disease caused by the parasite *Philasterides dicentrarchi*.

Abstracts

Celebrating 10 years of FAANG - From FAANG to Fork: Highly annotated genomes as resources to improve farmed animal production

Emily Clark, University of Edinburgh

Sustainable improvements in the efficient production of farmed animals will be needed in coming decades to provide healthy food for a rapidly growing human population. The challenge is to produce more food using fewer resources, in a sustainable way that meets societal expectations, and mitigates the effects of rapidly changing climates. Chief amongst the improvements required in farmed animal breeding is the ability to more accurately use an animal's genotype to predict it's phenotype. The Functional Annotation of Animal Genomes (FAANG) project is an international effort to characterise the functional elements of the genomes of farmed animals. Advances in the analysis of genome function will provide tools and knowledge to answer the genotype to phenotype question. The first 10 years of FAANG focused on foundational data generation to characterise expressed and regulatory genomic regions, curation and provision of highly annotated farmed animal genomes. These were largely based on individual level, high depth approaches. Knowledge gained from the FAANG projects is now being used by researchers across the globe to link genotype to phenotype in the major farmed animal species. Continued public investment, international collaboration, data infrastructure and training of new scientists are now needed to advance genotype to phenotype research further. In the FAANG to Fork research strategy we provide a framework of priorities for the next 10 years. The primary challenge facing this community now is harnessing the resources generated to link genotype, phenotype and genetic merit in order to translate this research out of the laboratory and into industry application in the field.

Abstracts

Evolution of duplicated genome regulation in salmonids

Marie-Odile Baudement, NMBU

After characterizing chromatin states of Atlantic salmon and rainbow trout, we focused on promoter and enhancer regulatory elements. Enrichment of TFBS in these regions revealed that active enhancer regions exhibited 141 common TFBS across species, and 46/53 species-specific TFBS. Conserved vertebrate TFBSs including POU5F1, Sox and Hox were identified for both species during developmental stages, alongside conserved TFBS specific to particular tissues, e.g. NRF1 for brain, PPARG for liver or MEF2B for muscle. These findings illustrate an overall conservation of TFBS usage for both salmonids. Clustering of expression of ohnolog-tetrads revealed five distinct categories of genes. A subset of genes fell into the category where both copies of the same species is expressed differentially compared to the copies of the other species. Moreover, a group of ohnolog-tetrads displayed similar expression patterns between both species, with either both copies being co-regulated or either only one copy in each species. We anticipate that gene expression differences are associated with differences in regulatory activity of chromatin. By exploring the gene expression profiles of the different Hox genes cluster in both Atlantic salmon and rainbow trout, we observed that despite the duplication of all the clusters, the transcription of each of the copies were maintained, with the exception of the HoxA-b copy, which appeared to be lost in rainbow trout, and is evidently on the path to pseudogenization in Atlantic salmon.

Abstracts

Gene expression evolution after the salmonid-specific autopolyploidization event

D. Perojil Morata[1], M.-O. Baudement[2], P. Dewari[1], G. Gillard[2], D. Baranasic[3,4], M.K. Gundappa[1], T. Podgorniak[2], L. Grønvold[2], A. Laurent[5], Aline Perquis[4], Erika Carrera Garcia[2], Thi Nguyen[2], R. Ruiz Daniels[1], G. Ilsley[6], P. Harrison[6], D. Thybert[6], J. Bobe[5], C. Berthelot[7,8], E. Parey[5,7], A. Louis[7], F. Giudicelli[7], H. Roest Crollius[7], T. Hvidsten[2], S. Sandve[2], B. Lenhard[3,4], Y. Guiguen[5], M. Kent[2], S. Lien[2], D.J. Macqueen[1]

1. The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, UK.
2. Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Centre for Integrative Genetics (CIGENE), Norwegian University of Life Sciences, Ås, Norway.
3. MRC London Institute of Medical Sciences, London, UK.
4. Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, London, UK.
5. INRAE, LPGP, Rennes, France.
6. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK.
7. Institut de Biologie de l'ENS (IBENS), Département de Biologie, École Normale Supérieure, CNRS, INSERM, Université PSL, Paris, France.
8. Institut Pasteur, Université Paris Cité, CNRS UMR 3525, INSERM UA12, Comparative Functional Genomics group, F-75015 Paris, France.

Whole genome duplication (WGD), where ploidy level is doubled in a single event, has been proposed as a key driver of diversification and the evolution of new morphological and physiological traits. The ancestor of salmonid fishes underwent a lineage-specific WGD event by autopolyploidization (i.e. spontaneous WGD) ~ 100 Mya, and salmonids today are still undergoing rediploidization, making them an ideal vertebrate system to study evolution after WGD. Past studies have investigated the impact of the salmonid-specific WGD on the retention, expression, and regulation of genes retained as duplicate copies (ohnologues). Nonetheless, to date, we have a limited understanding of how the duplicated genome evolved under constraints operating at different stages of salmonid ontogeny, for example contrasting embryonic development with adult tissues. We addressed this knowledge gap by investigating gene expression across ontogeny using matched gene expression (RNA-seq) samples from two species sharing the ancestral salmonid-specific WGD, Atlantic salmon (*Salmo salar*) and Rainbow trout (*Oncorhynchus mykiss*), for embryonic stages as well as a panel of tissues, generated as part of AQUA-FAANG WP1. This was followed by a cross-species analysis of orthologous transcriptome similarity based on the Jensen–Shannon divergence method. The results support previous studies suggesting the period of highest phenotypic and genomic constraint during ontogeny, known as the phylotypic period, is found towards the end of somitogenesis. This finding was mirrored by a similar analysis within each species, which revealed that ohnologue pairs show the highest conservation of expression level at the inferred phylotypic stage, suggesting selection on ohnologue expression is influenced by ontogenetic constraints operating more widely during evolution. Further analyses revealed a higher proportion of ohnologues co-expressed during the inferred phylotypic stage, with the highest divergence in ohnologue expression found in the earliest stage sampled, during the transition from cleavage to blastulation, when the transcriptome is dominated by maternally deposited mRNA. Overall, this analysis reveals a complex picture of interactions between ontogenetic constraints and ohnologue expression and regulation following a WGD event. Further research into the role of regulatory network complexity and dosage sensitivity in ohnologue conservation could increase our understanding of evolutionary outcomes following WGD, such as the evolution of novel morphology.

Abstracts

COMPARATIVE REGULOMICS GIVES INSIGHTS INTO THE CONSERVATION AND EVOLUTION OF REGULATORY ELEMENTS FOLLOWING WHOLE GENOME DUPLICATION IN SALMONIDS

Manu Kumar Gundappa[1], Diego Perojil Morata[1], MO Baudement[2], Pooran Dewari[1], Gareth Gillard[2], Damir Baranasic[3], Tomasz Podgorniak[2], Lars Grønvold[2], Audrey Laurent[4], Aline Perquis[4], Erika Carrera Garcia[2], Thi Nguyen[2], Rose Ruiz Daniels[1], Garth Ilsley [5], Peter Harrison[5], David Thybert[5], Julien Bobe[4], Cervin Guyomar[4], Thomas Desvignes[6], Camille Berthelot[7,8], Elise Parey[4,7], Alexandra Louis[7], Francois Giudicelli[7], Hugues Roest Crollius[7], Torgeir Hvidsten[2], Simen Sandve[2], Boris Lenhard[3], Yann Guiguen[5], Matthew Kent[2], Sigbjørn Lien[2], Daniel J. Macqueen[1]

1. The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, UK.
2. Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Centre for Integrative Genetics (CIGENE), Norwegian University of Life Sciences, Ås, Norway.
3. MRC London Institute of Medical Sciences, London, UK.
4. Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, London, UK.
5. INRAE, LPGP, Rennes, France.
6. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK.
7. Institut de Biologie de l'ENS (IBENS), Département de Biologie, École Normale Supérieure, CNRS, INSERM, Université PSL, Paris, France.
8. Institut Pasteur, Université Paris Cité, CNRS UMR 3525, INSERM UA12, Comparative Functional Genomics group, F-75015 Paris, France.

The common ancestor of salmonids underwent a lineage-specific WGD event ~100 million years ago and a large proportion of the genome is retained in duplicate, offering an ideal vertebrate system to understand the role of WGD in genome evolution. The huge commercial importance of these species to aquaculture further demands improved understanding of genome function and regulation, which is still poorly understood. In the current study, we make extensive use of the functional annotation data generated through the European AQUA-FAANG project, including 0.6 billion ATAC-Seq and 4 billion ChIP-Seq reads, to investigate duplicated regulatory elements in the genomes of Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*). The objective was to examine the conservation of regulatory element activity through ontogeny in both species. Duplicate-aware whole genome alignments including Atlantic salmon and rainbow trout were generated with Cactus to align the duplicated syntenic regions in both species. High-confidence ATAC-Seq peaks (open chromatin regions) representing multiple stages of embryogenesis, and six adult tissues at two stages of sexual maturation, were overlapped with the Cactus alignments. The coupling of sequence and regulatory element conservation in open chromatin regions was established with respect to duplicated and orthologous regions.

Cross-referencing open chromatin regions with genome alignments revealed open chromatin regions overlapping both duplicated sequences (duplicated peaks) retained from WGD increased across embryogenesis, being highest at the late somitogenesis stage, and was variable across adult tissue types, with brain showing the highest proportion. Reciprocally, we identified the lowest proportion of open chromatin regions in singleton sequences (singleton peaks) at the equivalent stage of development and tissues. Overlapping conserved non-coding elements (CNEs) of different evolutionary ages revealed ancient CNEs being hugely overrepresented in duplicated peaks compared to singleton peaks through ontogeny while recently evolved CNEs showed no marked difference in their abundance across duplicated and singleton peaks through ontogeny. Our results validate the use of genome alignment to understand the dynamics of regulatory element activity across the duplicated genomes of salmonids. Our results are broadly consistent with the hourglass model of development

(Duboule, 1994), suggesting highest evolutionary constraints on gene regulation during the phylotypic stage, a pattern previously observed across species, but not in relation to WGD. The results of this work, by revealing conserved regulatory elements linked to salmonid phenotypes, will support the uptake of functional genomic information into salmonid genetics and selective breeding approaches supporting sustainable and profitable salmonid aquaculture.

Abstracts

Linking divergence of salmonid gene expression to regulation

Gareth Gillard, NMBU

The salmonids Atlantic salmon and rainbow trout share a past salmonid-specific whole genome duplication event, leaving about half of genes with a duplicate pair. Duplicate ortholog genes between both species are termed ohnologs. Ohnologs may have remained with conserved gene function, but can also diverge in function (for example, a difference in tissue specificity) as selection pressure is relaxed from having a redundant copy. How ohnologs become diverged is something we look for more evidence of. Logically, changes to gene regulation could drive this. Therefore with the functional genomic data generated for salmon and trout, we investigated how ohnolog genes may have diverged in gene regulation by relating differences in the regulatory landscape, annotated from chromatin state activity and robust open chromatin regions, to the gene expression profiles across adult tissues. We looked to see if conservation or divergence in gene expression relates to the same change in enhancer element composition surrounding the gene transcription start sites. This would provide evidence that regulatory based mechanisms can be a driving force behind gene expression shifts and possible evolution after gene duplication events.

Abstracts

Differences in transcription initiation in early embryogenesis between Cyprinid species

Damir Baranašić, Ada Jimenez-Gonzalez, Bojan Žunar, Yavor Hadzhiev, Ferenc Müller, Boris Lenhard

By integrating diverse cellular signals, promoters coordinate the transcription of every gene in the genome. As such, they are crucial for every organism's proper development and homeostasis. However, given their critical role in health and disease, our understanding of how they are built and change through evolutionary time is sparse. In order to gain insight into the mechanisms that initiate transcription, it is essential to understand the evolution of promoter structures. To explore this, we conducted cap analysis of gene expression (CAGE) on the developmental timeline of the cyprinid fish common carp (*Cyprinus carpio carpio*). CAGE is a highly impartial method for detecting transcription initiation at a single nucleotide resolution. With its whole genome duplication that occurred approximately 12 million years ago, the common carp provides an ideal organism for studying the evolution of promoter architecture. Its genome is an allotetraploid with separated subgenomes, which enables us to track the co-evolution of promoters within the same species. By systematically comparing transcription initiation and promoter structures between the two subgenomes, we were able to explore how substitutions in the initiator dinucleotide govern differences in ohnolog gene expression. Finally, we compared the CAGE data of common carp with those of zebrafish, a closely related species that did not undergo a recent whole genome duplication, to determine the effects of whole genome duplication on promoter architecture and activity.

Abstracts

A comparative epigenomic resource for regulatory element annotations: mapping non-coding functional elements during *Cyprinus carpio* embryo development

Ada Jimenez-Gonzalez[1], Damir Baranasic[2], Annemiek Blasweiler[3], Bojan Zunar[4], Hendrik-Jan Megens[3], Geert Wiegertjes[3], Boris Lenhard[2], Ferenc Müller[1]

1. Institute of Cancer and Genomics Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK.
2. MRC Clinical Sciences Centre, Imperial College, London, UK
3. Aquaculture and Fisheries, Wageningen University, The Netherlands
4. Laboratory of Biochemistry, Department of Chemistry and Biochemistry, University of Zagreb, Croatia

Common carp (*Cyprinus carpio*) is an economically important ornamental and aquaculture species. Interestingly, carp is an allotetraploid fish emerged by hybridisation 13 million years ago conserving two well-defined and chromosomally separated sub-genomes. These genomic features provide an ideal comparative platform for developmental genomics of cyprinids, including zebrafish. Carp separated from zebrafish approximately 60 million years ago, which is an optimal divergence time for reliable detection of functional non-coding elements by sequence conservation analyses. We aimed to investigate the regulatory landscape and coordination of gene expression between the two sub-genomes and their specific regulatory divergence during carp early development and to compare it to that of zebrafish (e.g. DANIO-CODE). To generate this regulatory element atlas and to exploit carp as comparative genomics resource, we have investigated chromatin accessibility by ATAC-seq, transcription start sites at promoters by CAGE-seq and cis-regulatory element-associated histone modification marks (H3K27ac, H3K4me3 and H3K27me3) by ChIP-seq from six key developmental stages. We integrated these chromatin characterisation data with our RNA-seq datasets from 12 developmental stages from unfertilized eggs through to prehatch stages. We have mapped candidate regulatory elements of carp development genome wide and studied their divergence in the two sub-genomes. The epigenomic profiles generated from carp will also serve as comparative epigenomic resource for synteny anchoring-based mapping of epigenomic domain conservation such as H3K27ac marked superenhancers or H3K27me3 marked regulatory landscape. Our approach provides information on conserved regulatory elements and insights into the evolution of teleost cis-regulatory landscape upon genome duplication.

Abstracts

Genomic and cellular insights into antiviral responses and viral disease resistance in salmonid fishes

Thomas Clark, Shahmir Naseer, Manu Kumar Gundappa, Audrey Laurent, Aline Perquis, R. Taylor, L. Jouneau, Bertrand Collet, Daniel J. Macqueen, Samuel A. M. Martin and Pierre Boudinot

Antiviral innate immunity is orchestrated by the interferon system, which appeared in ancestors of jawed vertebrates. Interferon upregulation induces hundreds of interferon-stimulated-genes (ISGs) with effector or regulatory functions. We investigated the evolutionary diversification of ISG responses through comparison of two salmonid fishes, accounting for the impact of sequential whole genome duplications ancestral to teleosts and salmonids. By analysing the transcriptomic response of the IFN pathway in the head kidney of both rainbow trout and Atlantic salmon in parallel *in vitro* and *in vivo* experiments, we identified a large set of ISGs conserved in both species and cross-referenced them with zebrafish and human ISGs. Thus, generating annotated tables detailing orthology and functional responses in both salmonid species.
We also have previously described isogenic rainbow trout lines with contrasted resistance to Viral Hemorrhagic Septicemia Virus (VHSV) and to Flavobacterium psychrophilum (Fp), two key pathogens in European aquaculture. To better understand mechanisms behind genetic resistance, we undertook a global analysis of immune cell subsets within the head kidney of these two rainbow trout lines. Our single-cell RNA-seq data of unstimulated control fish produced 34 clusters that could be separated into 6 major cell lineages: B-cells, T-cells, granulocytes, monocytes, thrombocytes and erythrocytes. Further investigation revealed clear differential expression of gene subsets between the two isogenic lines, comprising a number of immune genes.

Abstracts

Comparative regulomics in flatfish: from turbot to the main farmed Pleuronectiformes

Juan Rubiolo, University of Santiago de Compostela

We aimed at understanding the relationship between gene regulatory networks in flatfish (Pleuronectiformes), a group showing remarkable adaptations to demersal lifestyle. We searched the sequences of enhancers identified in the turbot atlas of regulatory elements by ChIP- and ATAC-seq in the chromosome-level assembled genomes of the five most relevant commercial flatfish species available in Ensembl together with turbot, *Cynoglossus semilaevis*, *Hippoglossus hippoglossus*, *Paralychtis olivaceous*, and *Solea senegalensis*, using in addition the *Oryzias latipes* genome as a well-annotated Acanthoptherigian outgroup. We blasted the sequences of turbot enhancers, with lengths ~400 bp, against the selected flatfish genomes. Results were filtered according to the quality and length of the alignment, and those with length ≥ 350 bases were used for a preliminary analysis. Using this filter, we identified ~ 900 enhancers shared by the 6 species included in the study. Less stringent filters resulted in increased number of shared enhancers encompassing until ~10,000 when applying a filter length ≥ 200 bp. Based on this information, we confirmed the syntenic patterns previously observed with coding genes in the same species. Then, we inspected annotated coding genes surrounding the genomic regions around each enhancer (±100 kb). Genes in these regions were evaluated for functionally enriched GO terms to identify conserved regulatory pathways among flatfish. Using HOMER, we also identified the motifs present in the shared enhancers and explored the putative transcription factors recognizing these motifs. Functional enrichment was also performed for the identified transcription factors and the results compared with the previous enrichment analysis. Multiple genome alignment was performed using CACTUS, which will be used for revising the sets of conserved regulatory elements detected across species for further functional comparative studies associated with productive traits in flatfish and teleosts.

Abstracts

Dynamic gene expression and regulation during gilthead sea bream development

Sarropoulou E.[1], Tsigenopoulos C. S.[1], Kaitetzidou E. [1], Slavka Plovina I.[1], Mylonas C.C

1. Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC), Hellenic Centre for Marine Research (HCMR), Thalassocosmos, Gournes Pediados, Crete, Greece

The Gilthead sea bream (*Sparus aurata*) is an economically important fish species that is well-adapted to farming and changing environments. A few transcriptome analyses have been used to study the expression and the regulation of gene expression during gilthead sea bream development. Here we present the dynamics of gene expression of twelve developmental stages ranging from the 32-cell stage to the stage of pharyngula. Obtained expression patterns follow the developmental time course while stage-specific expressed genes are also detected. To further enrich the knowledge on embryonic development gene regulation at the epigenetics level has been carried out. Hence, small non-coding RNA sequencing was carried out for six key developmental stages. Amongst the several classes of sncRNAs, the most prominent and well-studied ones are microRNAs (miRNAs) which have been already shown to play a key role in development. Within the present study important known and unknown have been detected with putative regulatory roles of chief developmental genes. Therefore, the mRNA-miRNA expression patterns and their putative interactions will be discussed. Additionally, for the six key developmental stages, ATAC-Seq (Assay for transposase Accessible Chromatin using sequencing) and ChIP-seq (Chromatin ImmunoPrecipitation sequencing) were carried out and preliminary results will be presented.

Abstracts

Day Two, Thursday 12th October

Genome functional annotation of host defense response in gilthead sea bream (*Sparus aurata*) through chromatin accessibility and differential gene expression assays

Tsigenopoulos C. S., Radojicic J.[1], Papadopoulou A.[1,2], Papadogiannis V.[1] Katharios P.[1], Papadakis I.[1], Sarropoulou E.[1], Manousaki T.[1], [1]

1. Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC), Hellenic Centre for Marine Research (HCMR), Thalassocosmos, Gournes Pediados, Crete, Greece

2. School of Medicine, University of Crete, Greece

The gilthead seabream is a fish species of high economic importance for the Mediterranean aquaculture industry, and the present study assesses for the first time its immune response map on bacterial and viral stimuli on epigenetic and gene levels. Aiming to improve the species' production and welfare, the scientists, hatchery and production managers should work to reduce the incidence of bacterial and viral infections. In parallel, a more holistic and profound understanding of the underlying mechanisms of the immune response will help to prevent and treat these health problems.

The development of genome-wide functional annotation maps ('ImmunoMaps') that represent host defense responses of the immune system was implemented in the AquaFAANG project (www.aqua-faang.eu) by using ATAC and RNA sequencing. We made use of two distinct classes of disease agents, neutralized *Vibrio* and polyinosinic:polycytidylic acid (poly I:C), following protocols developed in the project and optimized for each species separately. The head kidney immunity activation was explored *in vivo* by injecting fish and *in vitro* in primary head kidney leukocyte cultures; RNA and ATAC libraries were constructed and sequenced, and bioinformatic analyses for the combined sequencing data from both approaches were performed using pipelines from the nf-core initiative. The differentially expressed genes were mapped onto the gilthead sea bream genome and revealed few hundreds of up-regulated and down-regulated genes for Vibrio and PIC stimulations as well as virus- and bacteria-specific genes for the *in vivo* and *in vitro* experimental conditions. Last, the association of the differentially accessible peaks with the differentially expressed genes allowed the assembly of a core immune related gene network, and unveiled a potential regulatory link between the up-regulated genes next to under-accessible peaks *in vivo* for all treatments (Vibrio-PIC, merged together) where an over-representation for interferon-gamma response terms was present.

Abstracts

TRANSCRIPTOME AND CHROMATIN LANDSCAPE OF EUROPEAN SEABASS IMMUNE RESPONSE TO VIRAL-LIKE STIMULATION

Serena Ferraresso, Rafaella Franch, Massimiliano Babbucci, Giulia Dalla Rovere, Luca Peruzza, Daniela Bertotto, Francesco Pascoli, Anna Toffan, Luca Bargelloni

Department of Comparative Biomedicine and Food Science, University of Padova, Italy
Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe), Italy

Disease outbreaks account for almost the 10% of the European seabass mortality in aquaculture farms. Knowledge of the complex molecular mechanisms underlying host response to pathogen is of crucial importance to improve understanding of genome function and regulation.
As part of the AQUA-FAANG (www.aqua-faang.eu) project, the present work aims to profile *in vivo* and *in vitro* response to stimulation with viral (Poly I:C) mimics.
In vitro challenge was conducted on head kidney isolated leucocytes stimulated with PBS or Poly I:C and collected 12 hours post-infection (hpi). For in vivo challenge, adult individuals were stimulated by injection with PBS or Poly:IC, animals were sacrificed 24 hpi and head kidney sampled for subsequent analyses. All samples were employed for RNA-seq, ATAC-seq and ChIP-seq library preparation. Differentially expressed genes (DEGs) analysis was conducted with EdgeR while ChromHMM was employed on ATAC-seq and ChIP-seq data in order to define genome-wide chromatin states
Differential expression analysis identified a total of 648 and 1491 DEGs from *in vitro* and *in vivo* datasets, respectively. For both experiments, Over-representation Analysis (ORA) and Gene Set Enrichment Analysis (GSEA) highlighted the significant enrichment of gene sets related to immune system and viral infection in fish, in particular Interferon responsive genes.
ChromHMM analyses allowed to define significant differences in chromatin states regions between conditions, with several immune-related pathways enriches in genes that acquired active chromatin states following Poly:IC stimulation.
The present study provides the first genome-wide functional annotation map of the European seabass response to immune stimulation

Abstracts

Functional genomic architecture of viral nervous necrosis disease resistance in farmed European seabass

Robert Mukiibi[1], Luca Peruzza[2], Carolina Penaloza[3], Massimiliano Babbucci[2], Matteo Fregugli[4], Stanis Laureau[4], Daniela Bertoto[2], Rafaella Franch[2], Giulia Dalla Rovere[2], Serena Ferraresso[2], [2]Sara Faggion, [5]Costas Tsigenopolous, Ross D Houston[3], Luca Bargelloni[2] and Diego Robledo[1]

1. The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Edinburgh, United Kingdom

2. Department of Comparative Biomedicine and Food Science, University of Padova, Italy

3. Benchmark Genetics, 1 Pioneer Building, Edinburgh, Technopole, Penicuik, United Kingdom

4. Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC), Hellenic Centre for Marine Research (H.C.M.R.), Crete, Greece

5. Valle Cà Zuliani Società Agricola s.r.l., Conselice (RA), Italy.

Viral nervous necrosis (VNN) caused by the nervous necrosis virus (NNV) is a major infectious disease threatening the European seabass aquaculture industry. VNN causes high economic losses emanating from high mortality rates and slow growth of infected fish. Selective breeding has the potential to increase the disease resistance of aquaculture stocks, reducing the impact of disease. Although recently numerous studies have identified a major QTL associated with VNN resistance the genes and genomic variants modulating resistance to the disease remain unknown. In the current project, we integrated multiple functional genomic tools including sequence-based GWAS, sequence-based targeted eQTL, ATAC-seq, and ChIP-seq analyses to identify causal variants and genes modulating VNN resistance in farmed European seabass. Like previous studies, our results demonstrated that VNN resistance is moderately heritable (h^2 = 0.45). Additionally, GWAS results confirmed a major VNN resistance QTL on LG12. Variants in this QTL region explained up to 38.3% of the additive genetic variance of resistance to the disease. Interestingly the most significant variants were located within the *IFI27L2A* gene. Furthermore, our analyses showed a remarkable association between *IFI27L2A* gene expression and VNN resistance in the brain and head kidney. Expression quantitative trait loci analyses further revealed that GWAS significant variants also had a significant impact on the expression of *IFI27L2A* in both brain and head kidney. Indeed, ATAC-seq analyses identified three overlapping active chromatin regions in the QTL/eQTL region containing the *IFI27L2A* in the brain and head kidney tissue.

Abstracts

Decoding enhancer function: from the nucleosome to the nucleus

Wendy Bickmore, MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh

The complexity of vertebrates comes not from the number of genes in the genome but from the 100s of thousands, perhaps millions, of enhancers that regulate gene expression with exquisite precision in time and space. Identifying enhancers, the genes they regulate, and the biological context in which they operate remains a major challenge. Moreover, we still do not understand how enhancers communicate to their target genes, often over long genomic distances.

I will discuss our efforts to drive progress in understanding when, where and how enhancers function in development using cellular and animal models, from the level of histone modifications through to the 3D structure of the genome in the nucleus.

Abstracts

Transcriptional differences in CyHV-3 response between resistant and susceptible common carp (Cyprinus carpio) crossings.

Lukasz Napora-Rutkowski[1], Luca Peruzza[2], Joanna Szczygieł[1], Teresa Kamińska-Gibas[1], Luca Bargelloni[2], Ilgiz Irnazarow[1]

1. Polish Academy of Sciences, Institute of Ichthyobiology and Aquaculture in Golysz (ZIGR), Poland
2. Department of Comparative Biomedicine and Food Science, University of Padova, Italy Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe), Italy

Common carp (*Cyprinus carpio*) is one of the five most produced aquaculture species in EU countries, where Poland and Czechia are the leading breeders of this fish. One of major threat to common carp aquaculture is still Cyprinid herpesvirus-3 (CyHV-3) infection, leading to widespread mortality and substantial economic loss. Anti-disease selective breeding strategy is very promising in obtaining common carp lines with high level of CyHV-3 resistance. Amur carp (*C. c. haematopterus*) originated from the Amur River basin have a high resistance to CyHV-3 virus infection. This fish were selected for crossbreeding with polish common carp breeding lines to produce CyHV-3 resistant families. Two of most resistant (R1:82% ±6.8 and R2:66% ±10 survival) and two most susceptible backcross families (S1:16% ±5 and S2:1% ±1.7 survival) were challenged with CyHV-3 virus. Tissues of most important immunological fish organs: head-kidney, liver, and spleen were sampled pre-infection and at day of highest mortality (11) for RNA-Seq analysis. Transcriptome differential expression analyses was performed between the resistant and susceptible groups of backcross families, followed by pathway enrichment analysis. Two types of functional analyses were performed: Over-representation analysis (ORA) and Gene-Set Enrichment Analysis (GSEA). This study present the functional pathways and immune mechanisms involved in both sensitive and resistant families during CyHV3 infection in common carp.

Abstracts

Symmetric expression of ohnologs encoding conserved antiviral responses in tetraploid common carp suggest absence of subgenome dominance after whole genome duplication

Blasweiler[1] A., Megens[2] H-J., Goldman[1] M.R.G., Tadmor-Levi[3] R., Lighten[4] J., Groenen[2] M.A.M., Dirks[5] R.P., Jansen[5] H.J., Spaink[6] H.P., David[3] L., Boudinot[7] P., and Wiegertjes[1] G.F

1. Aquaculture and Fisheries, Wageningen University, The Netherlands
2. Animal Breeding and Genomics, Wageningen University, The Netherlands
3. Dept. of Animal Sciences, RH Smith Faculty of Agriculture Food and Environment, The Hebrew University of Jerusalem, Israel
4. Biosciences, University of Exeter, United Kingdom
5. Future Genomics Technologies B.V., The Netherlands
6. Institute of Biology, Leiden University, Leiden, Netherlands
7. Université Paris-Saclay, INRAE, UVSQ, VIM, Jouy-en-Josas, 78350, France.

Allopolyploids often experience subgenome dominance, with one subgenome showing higher levels of gene expression and greater gene retention. Here, we address the functionality of both subgenomes of allotetraploid common carp (*Cyprinus carpio*) by analysing a functional network of interferon-stimulated genes (ISGs) crucial in anti-viral immune defence. As an indicator of sub-genome dominance we investigated retainment of a core set of ohnologous ISGs. To facilitate our functional genomic analysis a high quality genome was assembled (WagV4.0). Transcriptome data from an *in vitro* experiment mimicking a viral infection was used to infer ISG expression. Transcriptome analysis confirmed induction of 88 ISG ohnologs on both subgenomes. In both control and infected states, average expression of ISG ohnologs was comparable between the two subgenomes. Also, the highest expressing and most inducible gene copies of an ohnolog pair could be derived from either subgenome. We found no strong evidence of subgenome dominance for common carp.

Abstracts

Multiomics reveals the genomic regulatory landscape underlying the antiviral response in Atlantic salmon

S. NASEER[1], T. C. CLARK[2], B. COLLET[2], P. DEWARI[3], D.J. MACQUEEN[3], P. BOUDINOT[2] AND S.A.M. MARTIN[1]

1. SCOTTISH FISH IMMUNOLOGY RESEARCH CENTRE, SCHOOL OF BIOLOGICAL SCIENCES, UNIVERSITY OF ABERDEEN, UK
2. UNIVERSITE PARIS-SACLAY, INRAE, UVSQ, VIM, JOUY-EN-JOSAS, FRANCE
3. ROSLIN INSTITUTE, UNIVERSITY OF EDINBURGH, UK

Examining the genome-wide regulatory response of Atlantic salmon to viral infection is central to understanding the control of antiviral immunity. This study investigated the epigenomic and transcriptomic response to stimulation with the viral mimic poly I:C in Atlantic salmon. We employed ATAC-seq and ChIP-seq in combination with RNA-seq analysis to comprehensively examine changes in chromatin accessibility, histone modifications and gene expression. We identified a core set of 197 interferon-stimulated genes (ISGs) that were epigenetically modulated and highly up regulated in response to poly I:C. Fifty-four of these genes were also ISGs in rainbow trout, zebrafish, and humans, highlighting their evolutionary conservation. Our analysis revealed key transcription factors involved in the interferon response, including IRF9, STAT1, and STAT2. Regulatory elements showing increased activity to poly I:C were enriched in conserved vertebrate binding sites for STAT6, PRDM1, IRF6, JDP2, NR2E1, and BCL6, suggesting their central roles in the antiviral immune response. Focused analysis of Interferon Stimulating Response Elements (ISRE) indicated that genomic regions containing ISREs were modulated in response to poly I:C stimulation. Finally, we investigated differences in response to poly I:C stimulation among key ISGs retained as salmonid-specific paralogues, including MX, RSAD2, IRF9, DHX58, STAT1, IRF7 and CD9. This revealed paralogue-specific enrichment of ISRE motifs in promoter regions. Overall, this study provides novel insights into the regulatory mechanisms underlying the antiviral response in Atlantic salmon and underscores the significance of the epigenetic landscape in transcriptional regulation.

Abstracts

In vitro mutant models for functional characterization of genes of the type I IFN pathway in salmonids

Thomas Clark[1,2], Bertrand Collet[1], Samuel A. M. Martin[2], Pierre Boudinot[1]

1. Université Paris-Saclay, INRAE, UVSQ, VIM, Jouy-en-Josas, 78350, France.
2. Scottish Fish Immunology Research Centre, School of Biological Sciences, University of Aberdeen, Aberdeen, AB24, 2TZ, UK.

While *in vivo* KO models are generated in zebrafish, they are expensive difficult to implement in aquaculture species. For many intrinsic and innate immune mechanisms, *in vitro* models in immortal cell lines constitute relevant systems. A successful strategy of genome edition has been established, based on the Chinook salmon (*Oncorhynchus tshawytscha*) embryo (CHSE) cell line that was previously modified to stably express both a monomeric enhanced green fluorescence protein (mEGFP) and Cas9 (CHSE-EC). Two single guide RNAs (sgRNAs) located in the first 50-nt of the ORF were typically produced for each gene of interest. For transfection, sgRNAs were mixed with the sgRNA targeting mEGFP and used to transfect target cells using the Neon electroporation system. Transfected cells were expanded, and nonfluorescent single-trypsinized cells were single-cell cloned by FACS or by limiting dilution. Genomic DNA of selected clones was PCR amplified and directly sequenced for validation of mutation; clones which did not express GFP were all mutated (null mutation, in both haplotypes), and 50 to 70% of these clones also showed mutations in the gene of interest. We used this strategy to generate a repertoire of KO cell lines in which key genes of the type I IFN pathways are mutated. Targeted genes included effectors such as *pkr and viperin* as well as sensors, signaling molecules and transcription factors like *stat1 and stat2, irf3, and irf7.* Our cell lines provide a new opportunity to test directly the importance of these genes for the innate antiviral response in teleost fish.

Abstracts

Multiomics uncovers the epigenomic and transcriptomic response to viral and bacterial stimulation in turbot.

Oscar Aramburu[1], Belén Gómez-Pardo[1], Paula Rodríguez-Villamayor[1], Andrés Blanco-Hortas[1], Jesús Lamas[1], Pooran Dewari[2], Diego Perojil-Morata[2], Daniel J. Macqueen[2], Carmen Bouza[1], Paulino Martínez[1]

1. University of Santiago de Compostela, Spain

2. The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, UK

This study establishes the first regulatory annotation of the innate immune response in the genome of turbot (*Scophthalmus maximus*), a farmed flatfish. Integrating RNA-Seq, ATAC-Seq and ChIP-Seq, we identified differentially expressed genes (DEGs) and regulatory regions showing altered activity in the primary hematopoietic organ (head kidney) following viral (Poly (I:C)) and bacterial (inactivated *Vibrio*) stimulation both *in vivo* and using primary extracted leukocytes. We identified 8,797 DEGs across conditions, with enrichment among the upregulated genes in transcriptional activation (*Vibrio*) and immune response pathways, e.g. IFN-gamma (Poly (I:C)), alongside downregulation of metabolic and cell cycle genes. Genome-wide chromatin states were predicted separately for the *in vitro* and *in vivo* stimulations using ATAC-Seq and ChIP-Seq with the histone marks H3K4me3, H3K27ac and H3K27me3. We identified ignificant differences in chromatin accessibility (20,617 *in vitro*, 59,892 *in vivo*) and H3K4me3-binded (11,454 *in vitro*, 10,275 *in vivo*) regions were observed between immune states. Overlapping DEGs with promoters showing differential accessibility or histone mark binding revealed a coupling of gene expression and chromatin state. DEGs with activated promoters were enriched for similar functions to the global DEG set, but not in all cases, suggesting some key regulatory genes were in a poised or a bivalent state. We identified transcription factor binding sites (TFBS) enriched in differentially activated promoters and putative enhancers, revealing many were common to viral and bacterial responses. Finally, we provide an in-depth view of the chromatin state variance in promoter regions of key TF-coding DEGs with differential activation signals between stimulations. This multi-omic investigation of the immune response advances our understanding of the epigenomic basis for teleost host defence mechanisms and provides novel genomic regions to target in future precision breeding applications.

Abstracts

Day Three, Friday 13th October

Ensembl gene annotation, regulation and variant effect prediction for aquaculture

Peter W. Harrison[1] and Ensembl Project team[1]

1. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, CB10 1SD, Cambridge, UK

Ensembl (https://www.ensembl.org) is a freely available platform for exploring sequences and genome annotations across the tree of life, producing high-quality genomic resources and tools for use in academic research and industry. For more than twenty years, Ensembl has developed infrastructure to deliver reference genome assemblies from public archives for the genomic interpretation of genes, regulatory regions, variants and comparative data. Ensembl is recognised as a Global Core Biodata Resource and ELIXIR Core Data Resource highlighting the fundamental role it plays in life sciences research for open access to high-quality annotated genomic data. Through our close collaboration as part of the AQUA-FAANG project, Ensembl has taken the wealth of data generated by the project to substantially improve or release new reference annotations for the six most commercially important European aquaculture species (https://projects.ensembl.org/aqua-faang/). This talk will provide an overview of the freely available Ensembl services and tools for gene annotation, regulation and variant effect prediction, with a particular focus on their application for the aquaculture industry. This will include our rich annotation references, homology through our new Cactus alignments, the first Ensembl regulatory builds outside of human and mouse in these key aquaculture species, and the investigative power of the Variant Effect Predictor tool that assists researchers and industry specialists with the annotation and prioritisation of genomic variants. This continues Ensembl's mission to help scientists and the aquaculture industry to identify genes that are involved in important traits for production and resilience to disease.

Abstracts

Accounting for overlapping annotations as biological priors in genomic prediction models of complex traits

Andrea Rau, INRAe/GeneSwitch

It is now widespread in farm animal and plant breeding to use genotyping data to predict phenotypes with genomic prediction models. Functional genomic annotations (e.g., the accessibility of chromatin or methylation status in relevant tissues), have the potential to provide valuable insight into the location and effect size of causal genetic variants underlying complex traits. Developing and validating genomic prediction models able to fully leverage such complex functional annotations for improved accuracy and interpretability was one of the aims of the H2020 GENE-SWitCH project. To this end, we defined and implemented a flexible framework for genomic prediction called BayesRCO to simultaneously take advantage of the availability of multiple functional genomic annotations. In this talk, I'll describe the intuition behind our proposed model and discuss some of our key take-away messages from early results.

Abstracts

Functional genomics and selective breeding in aquaculture: implications from the AQUA-FAANG project

S. Faggion[1], R. Mukiibi[2], L. Peruzza[1], D. Bertotto[1], M. Babbucci[1], R. Franch[1], G. Dalla Rovere[1], S. Ferraresso[1], C. Peñaloza[2,3], R. Houston[2,3], P. Carnier[1], D. Robledo[2], L. Bargelloni[1]

1. Department of Comparative Biomedicine and Food Science, University of Padova, Italy

2. The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Edinburgh, United Kingdom

3. Benchmark Genetics, Edinburgh, United Kingdom

Selective breeding to enhance disease resistance is considered a feasible approach to prevent and control mortality from disease outbreaks. Due to the complexity in obtaining phenotypes for resistance to diseases, traditional selective breeding based on the routinary estimation of the genetic merit of the animals is not a feasible option. For this reason, the implementation of genomic selection procedures is of great interest, but improved ability and accuracy in using the genotype to predict the phenotype are required.

In AQUA-FAANG, a wide amount of functional genomic data has been produced, expressed and regulatory genomic regions have been characterized, and annotated genomes for aquaculture species have been provided. The integration of these data into genomic prediction models for complex traits such as disease resistance is a further step towards precision breeding.

In work package 5 of the AQUA-FAANG project, regulatory or functional elements (i.e. open chromatin regions) in the genome of European sea bass (*Dicentrarchus labrax* L.) were detected through ATAC-seq analyses; these information were used to filter the whole-genome sequences (> 6 million SNPs). The obtained SNP subsets were used to predict genomic estimated breeding values (GEBV) for viral nervous necrosis (VNN) resistance, with moderate accuracies (0.33-0.35). Further criteria (Chip-seq data, eQTLs, chromatin status scores, Ensembl regulatory data) to prioritize SNPs in genomic prediction models will be tested soon.

Genome-wide association analyses using whole-genome sequenced animals revealed a major QTL associated with VNN resistance phenotypes on linkage group 12. A total of 528 and 578 SNP markers were identified as significantly (FDR < 0.05) associated with VNN resistance as a binary trait and time to death, respectively. Using evidence from eQTLs and chromatin accessibility data, a putative causal variant was identified. The obtained information could be useful in designing a SNP array with a specific marker for VNN resistance.

Abstracts

Developing a flexible, low cost, multifunctional genotyping solution for selective breeding in aquaculture

Rachael Wilbourn, Xelect

Selective breeding in Aquaculture is moving away from classical pedigree-based genetic improvement, to using more advanced genome-wide SNP genotypes to better understand genetic relationships among individuals; improve selection accuracies and increase genetic gain sustainably. These tools are particularly relevant to traits that cannot be directly measured on selection candidates such as disease resistance; performance under production conditions or destructive traits (e.g. fillet yield). While there many competing genotyping technologies with a wide range of costs and capabilities, genomic selection is becoming much more affordable and attractive to farmers. Xelect, a Scottish-based specialist Aquaculture genetics provider, is already at the forefront of implementing pedigree and genomic-based selective breeding programmes to finfish and shellfish producers worldwide, using cost-effective SNP genotyping tools. The incorporation of functional genomic information into genomic prediction models offers the opportunity to significantly improve prediction accuracy for key commercial traits. As part of the AquaFaang project, Xelect developed flexible and affordable low density genomic tools that were enriched for SNPs significantly associated with functional variation for disease resistance traits in two key European Aquaculture species; European seabass (*Dicentrarchus labrax*) and Gilthead seabream (*Sparus aurata*). These two panels were validated using samples representing the range of genetic variation within the species and are now available for routine use by industry stakeholders, to improve our understanding of functional variation in economically significant traits, and determine their potential to be used to improve genomic prediction accuracy in commercial breeding programmes.

Abstracts

The way out of the black-box genomic selection

Kause Antti

Aquaculture breeding programmes typically use genomic evaluations that are based on genomic relationship matrix calculated based on thousands of SNP markers. This is a successful and powerful method but still a black-box approach in which minimal knowledge about the genomic determination of traits is needed. As a futuristic vision, it can be stated that the ultimate goal is that in the distant future, breeding value evaluations are based on the known genes. Both AquaIMPACT and AquaFaang EU-projects are in fact working towards this goal, yet we are at the very initial steps of this process. In AquaIMPACT project, we have shown that the accuracy of genomic breeding values can be increased by specifically modelling the mode of inheritance of traits (e.g. QTLs), and that this is a useful approach especially for disease resistance traits. The goal of AquaFaang is similar, to develop methods for more accurate precision breeding. The industry relevant applications of the Aqua-FAANG approach that can improve genomic evaluations are, a) small but precise SNP panels that have very low genotyping costs, b) SNP panels in which markers are closer to the actual genes, making genomic prediction more accurate, and c) identifying genes determining fish traits that can be especially modelled in genomic breeding value evaluations, and d) reveal currently hidden genetic variation that cannot be exploited by the current SNP panels. All these are steps towards more precise and efficient breeding in which the increased understanding of the genome is a key. Yet, more advanced statistical models of genomic evaluations are needed to fully exploit the accumulating new genomic information.

Abstracts

Advancing selective breeding in aquaculture through the functional annotation of fish genomes

Mark Looseley, Xelect

Commercial aquaculture production represents a diverse range of species; genomes and reproductive biology, requiring a corresponding diversity in breeding strategies and technologies. The use of genetic and genomic technology in aquaculture breeding has traditionally lagged behind its application in terrestrial livestock breeding. Nevertheless, high genetic diversity along with a large number of decentralised breeding programmes offer scope for significant genetic progress in key commercial traits if such technology were to be more widely adopted. Pathogens in particular, cause substantial losses to aquaculture producers as well as causing significant animal welfare concerns. The AQUA-FAANG project has produced new protocols and datasets for the functional annotation of fish genomes, leading to improved genome assemblies and functional annotations for key European aquaculture species in general, and for disease resistance loci in particular. Each of these outcomes has the potential to inform the development of precision breeding practices in aquaculture, increasing rates of genetic gain and improving the profitability and sustainability of production. The generation of routine genetic tools from these research outputs, along with accessible analytical tools and detailed modelling of potential economic returns will be critical to generating commercial impact from the substantial improvements in our understanding of genome function generated under the AQUA-FAANG project.

Abstracts

Regulatory Elements in Genomic Selection

Tim Knutsen, Aquagen

AquaGen, a research-oriented breeding company, plays a vital role in the advancement of the global aquaculture through the development and delivery of genetically optimized starter material and fertilized eggs. As an industry partner in the Aqua-FAANG project, this presentation will detail how AquaGen aims to utilize the wealth of functional genomic information generated by the Aqua-FAANG project to possibly enhance our breeding programs.

Specifically, the identification and utility of activated regulatory elements in the head kidney by immune stimulants, as well as those believed to influence sexual maturation, will be explored. Assessments both within and across generations will be conducted to evaluate the accuracy of genomic selection models that incorporate genomic markers within these regulatory regions. Additionally, the potential implications of 'functionally enriched' SNP-arrays for precision breeding will be discussed.

Abstracts

Potential use of AQUA-FAANG results to develop different breeding programs in diverse aquaculture companies increasing their efficiency, profitability and sustainability.

Adrián Millán, Geneaqua SL

In order to improve their production in a profitable and sustainable way, different aquaculture companies develop breeding programs with different objectives, especially focused on increasing growth, disease resistance and other complex traits such as fecundity or carcass quality. To improve complex traits, genomic selection is used, which requires genotyping and phenotyping thousands of individuals with panels of thousands of SNPs, associated with a very high cost and a complex operation to develop in production plants (disease challenge test, for example).

This talk will reflect from a genetic point of view on the needs and objectives of different aquaculture companies, indicating the possibility of implementing the knowledge derived from AQUA-FAANG to improve their efficiency. The functional genomic analysis of different tissues from different species and stages (BodyMaps-DevMaps), the knowledge of the regulatory regions involved in the response to viruses or bacteria (ImmunoMaps) as well as the discovery of causal variants associated to different QTLs will allow to advance in the selection. The use of low-density SNP panels incorporating functional information will reduce the cost of genotyping, the number of individuals to genotype and phenotype and facilitate the extrapolation of results between different populations analyzed, which is very important for the aquaculture industry in terms of economics, sustainability and animal welfare.