Advancing European Aquaculture by Genome Functional Annotation

Project no:	817923
Call:	H2020-SFS-2018-2
Start date:	1st May 2019
Duration:	48 months
Coordinator:	NMBU

D1.1 Development of standardized, optimized and processing efficient lab protocols

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923

Deliverable Name	Development of standardized, optimized and processing efficient lab protocols		
Deliverable No	D1.1		
Work package number(s)	WP1		
Document type (nature)	Report		
Due Date	31 Jan 2020		
Responsible Partner	NMBU		
Author(s) ***Name and Organisation***	Matthew Kent, NMBU Dan Macqueen, UEDIN		
Reviewer(s)	Sam Martin, UNIABDN Ross Houston, EDIN		
Dissemination level	PU	Public	X
	CO	Confidential, only for members of the consortium (including the Commission Services)	
Short description	Description of core lab protocols to be used across the project for DevMap and BodyMap sampling and library preparation.		

Change Records			
Version	Date	Changes	Author

Disclaimer

Content

1 Executive summary 3

2 Insert Header ..... **Error! Bookmark not defined.**

3 Annexes 9

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923

1 Executive summary

A key outcome from WP1 is sequencing data generated from ATAC-Seq, ChIP-Seq and RNA-Seq libraries that can be used for functional annotation of the six AQUA-FAANG species.

Because the project deliverables demand reliable comparisons of data across these species, it is important to standardize the protocols as much as possible so that each lab can generate comparable, high-quality data. But it is also critical to understand and emphasise that there is a high level of detail-complexity arising from the fact that there is a wide sample diversity (i.e. embryonic vs adult individuals, different tissues and species). This deliverable (D1.1) therefore describes the current standard operating procedures (SOP) for i) sample collection and library preparation developed using biological material from mature Atlantic salmon (*Salmo salar*) and ii) the sampling of salmonid embryos, ready for downstream library preparation. It is a detailed guideline that will enable partners to complete lab processing in a standardized manner, but deviations are expected across labs due to differences in infrastructure, species-specific tissue type, and sampling stage. We expect these deviations to be relatively minor, and that providing a suite of understandable SOPs will allow the project to meet its objective of uniform production of high-quality sequencing libraries across laboratories.

2 Body-Map vs. Dev-Map

Task 1.1 seeks to develop standardized functional annotation protocols that work using input from a variety of tissues harvested from mature and immature fish (BodyMaps) as well as from whole embryos at different developmental stages (DevMap). These are very different biological inputs which demand different approaches for sampling and library preparation. Mature and immature fish must be killed, dissected, organs identified and frozen, while embryos must be developmentally staged, the chorion and main constituents of the yolk removed, then preserved using different approaches. For this reason we have developed distinct protocols suited to BodyMaps and DevMaps; these are general protocols (available here) have been developed by testing real samples from Atlantic salmon, and as mentioned above, partners may/will need adjust aspects of the protocol to achieve optimal results.

3 Body-Map protocols

For BodyMap protocols, considerable testing has been performed to account for different properties of the target tissues (e.g. firm gill vs. pliable liver), and tissue specific guidelines are included at appropriate points in the protocols.

3.1 Sampling

Detailed protocols will be made available and described in Deliverable 1.2

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923

3.2 ATAC-Seq protocol

The ATAC-Seq protocol available here includes a detailed description of the following steps starting from frozen tissue:

(i) physical disruption of tissue and cell structure, release of intact nuclei, and dissolution of mitochondria,
(ii) nuclei isolation via centrifugation with subsequent manual counting,
(iii) introduction of Illumina sequencing adapters using transposase,
(iv) amplification of transposed DNA,
(v) size selection and quality control of library prior to sequencing

NOTES:

- In our hands, the protocol has been relatively robust, but we propose a few minor adjustments to account for different tissues (for example, additional tissue mincing when starting with gill, additional filtration when using skeletal muscle, a more gentle centrifugation to collect nuclei from head-kidney etc). We expect that optimizations will be required when beginning with tissue from immature (especially for processing gonad) and different species. NMBU will assist with guidance and advice, but partner labs will need to explore specific issues themselves and provide modified protocols that will (when finalized) be made publicly available.
- Adaptations will also be required when input is not solid tissue, but instead is cryopreserved dissociated cell mass from DevMap sampling (WP1) or primary cell cultures (WP3). Changes will likely be restricted to very initial ATAC-Seq stages; for example cryopreserved cells will need to be rapidly thawed, washed to remove preservative, and resuspended in homogenisation buffer. After this step, we envisage that processing will follow the prescribed general protocol.
- The protocol we release here is based on existing published protocols offering a linear series of steps. Our testing has highlighted which of these is critical/problematic (for example the number of strokes during dounce homogenization, which is influenced both by tissue and strength and style of the person performing this) as well as checkpoints where the successful execution of the steps can be assessed. We provide recommendations about how to judge a successful library preparation, specifically the observation of a nucleosome banding pattern.

3.3 ChIP-Seq protocol

A general ChIP protocol is available here, but particularly sensitive steps will need to be optimized by project partners to account for differences between tissues, species and lab infrastructure. It describes the following steps starting from frozen tissue:

(i) physical disruption of tissue and cell structure, release of intact nuclei,
(ii) crosslinking and counting of nuclei,
(iii) nuclei lysis and sonication
(iv) precleaning chromatin, antibody loading to beads, immunoprecipitation
(v) de-crosslinking, DNA purification, qPCR to assess enrichment
(vi) library preparation

NOTES:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923

- Cross-linking is a time sensitive step that will require optimization by each partner lab.
- Sonication is critical to generate short fragments suitable for subsequent capture and sequencing. In our lab we use a probe sonicator and fragment DNA according to specific power settings and exposure times (which are detailed in the shared protocol). Other partner labs will use different instrumentation for sonication for which our guidelines will not directly translate. They will, therefore need to empirically test sonication conditions and compare fragment distributions to determine the appropriate settings.
- A later step in ChIP involves performing a qPCR reaction to assess the relative enrichment of target DNA after immunoprecipitation. Targets for qPCR must be decided for each species and specific primers ordered and tested for their suitability.

3.4 RNA-Seq

The consortium has agreed that it is not important for all partners to use the same approach to store biological for RNA-Seq nor to extract total RNA from these samples, provided the methods used maintain all RNA types (i.e. small RNA, mRNA and long-non coding RNA). Samples can be processed by at least three options: i) fixation in RNAlater or equivalent buffer (following manufacturers guidelines) with subsequent storage at -80°C, ii) homogenized in Trizol or equivalent buffer (following manufacturers guidelines) with subsequent storage at -80°C, or iii) flash freezing using dry ice/ liquid nitrogen with subsequent storage at -80°C. The critical consideration is that extraction of each sample (i.e. biological replicate) must yield at least 2 µg (and ideally >5µg) total RNA of sufficient purity (judged by spectrophotometric 260/280 and 260/230 ratios) and integrity (judged by RNA integrity number, e.g. measured by Agilent BioAnalyser or Tapestation)

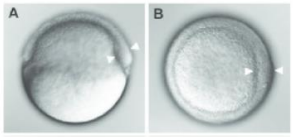
4 DevMap protocols

4.1 Selection of DevMap samples

Six key developmental stages are prioritized for sampling to generate ATAC-Seq, ChIP-Seq, and RNA-Seq libraries in each of the six species.

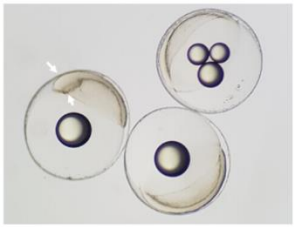
i) Late blastula. This will capture a stage soon after zygotic genomic activation, which begins earlier in the mid-blastula.

ii) Mid-gastrula. The key landmark targeted is the embryonic shield stage, where there is a slight thickening at the leading edge of the blastoderm during the formation of the germ ring. See below examples.

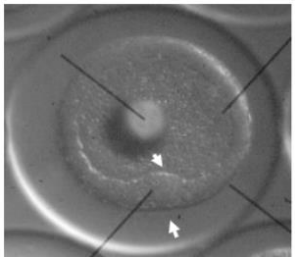


Zebrafish – organizer/shield highlighted

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923



Seabass – organizer/shield highlighted



Seabream – organizer/shield highlighted

iii) Early somitogenesis. A stage equivalent to the first 10% of the final somite number (e.g. if the final somite number is 60 – this would be between somite number 1 and 6)

iv) Mid-somitogenesis. A stage where the number of somites equates to half the final somite number +/- 5% the final somite number (e.g. if the final somite number is 60 – this would be between somite number 27 and 33)

v) Late somitogenesis. A stage equivalent to the last 10% of the final somite number (e.g. if the final somite number is 60 – this would be between somite number 54 and 60). Most rostral somites are not yet chevron shaped, the otic vesicle is formed.

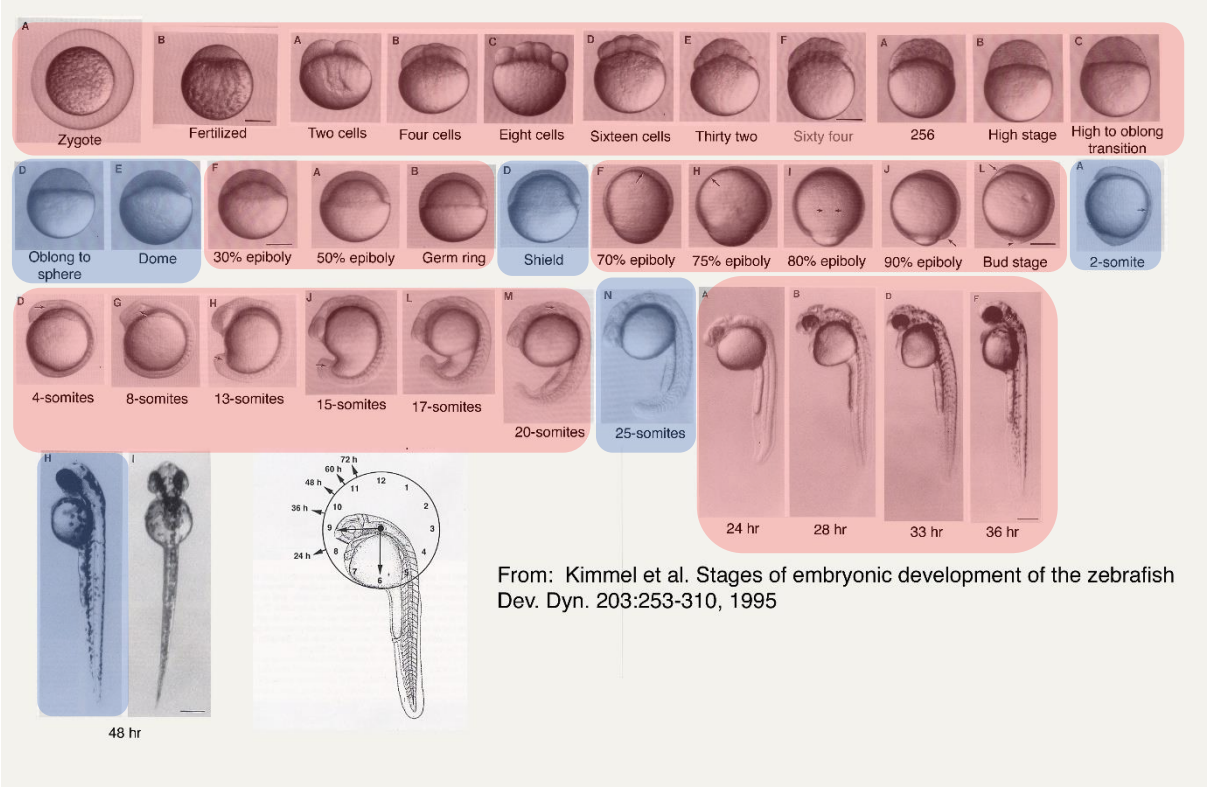
iv) Post segmentation - hatching or near hatching, eye pigmented in some species (e.g. salmon and zebrafish [= long pec stage]), fin formation started, organogenesis advanced.

In addition, additional regular sampling for total RNA will be done from before fertilization, through cleavages, blastulation, gastrulation, somitogenesis, and post-segmentation to hatching. The aim of this approach is to capture a higher resolution picture of development than the 6 key stages, which can be used for comparisons of RNA expression. The specific number of samples to take additional to the key stages will depend to some extent on the constraints posed by different species (e.g. some develop very fast), but the overall goal is to achieve a broader bank of samples available for RNA-Seq.

See below schematic indicating the Targeted Stages mapped against a staging criteria for zebrafish (*Danio rerio*):

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923

Blue shading – key stages targeted across species (mapped onto zebrafish development)
Red shading – *indicative* stages to capture in routine RNA preservation – the aim being to get multiple representative stages during cleavage, blastulation, gastrulation, somitogenesis and post-segmentation



4.2. Sampling protocols for ATAC-Seq

Trials with different species have indicated that flash freezing is not a safe option for preserving embryonic material for ATAC-Seq, as the nuclei are damaged. Consequently, a different approach was needed than for BodyMaps.

Two SOPs have been developed to preserve fish embryonic cells/nuclei for ATAC-Seq:

i) "Cell dissociation and slow freezing for ATAC-Seq using salmonid embryos" (available here)

ii) "Nuclear isolation from zebrafish embryos" (available here)

It was agreed among DevMap participants that as long as clean intact nuclei can be isolated from embryonic material, then either of the above SOPs should work well for ATAC-Seq and are suitable for use in the project. Additionally, other SOPs that isolate clean intact nuclei will be suitable for use in the project.

ATAC-Seq library preparations will adapt the protocol given above for BodyMaps (Section 3.2) using clean nuclei as the starting material, which can be obtained using the above SOPs.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923

4.3. Sampling protocols for ChIP-Seq

To maintain alignment with BodyMaps, all embryonic material used for ChIP-Seq in the project will be cross-linked.

The SOP used for preserving samples in Atlantic salmon is given below:

"Cross-linking DNA ready for ChIP-Seq using salmonid embryos" (available here)

This SOP should be suitable across different species, as it is generic. Note the guidance for different fixation times dependent on stage sampled. The objective is to sample 1 million cells per biological replicate, which will be dependent on the stage sampled and species (e.g. late blastula will have 3,000 to 10,000 cells across the AQUA-FAANG species). This goal may be challenging for the early stages in some species.

ChIP-Seq library preparations will adapt the protocol given above for BodyMaps (Section 3.3) using the cross-linked material as the input.

4.4. Sampling protocols for RNA-Seq

Please refer to Section 3.4, where the provided information applies equally to DevMaps. See the below SOP as an example of the selection RNA preservation method for Atlantic salmon:

"Total RNA preservation from salmonid embryos using Trizol" (available here)

4.5. Imaging developmental stages

The consortium agreed that it is important to have a standardized set of images for all developmental stages used in functional genomic studies spanning the different species. This is required to evidence that all samples used in genomics experiments match to the targeted developmental stages. A challenge is that not all species can be live-imaged, e.g. salmonid embryos are not transparent and fixed material works better for imaging than fresh material; however, other species are compatible with achieving high quality images of live embryos during sampling. Another challenge is that different partners have different imaging microscope systems.

With these constraints in mind, it recommended to:

i) take multi-angle images of every sampled developmental stage across all species using the equipment available in each lab
ii) fix 10-50 embryos at every sampled stage (see "SOP for fixation of salmonid embryos for imaging and staging**",** available here) which can be used for retrospective imaging analyses.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923

5 Future protocol versions

There is a very high probability that the protocols we make publicly available on the AQUA-FAANG website now will be modified and refined in the future. This will reflect expected workflow improvements and customisation to specific species and/or tissues. New versions will be dated and the specific modifications that have been made, highlighted.

6 Annexes

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817923