ESF Research Networking Programme

Conservation Genomics: amalgamation of conservation genetics and ecological and evolutionary genomics (ConGenOmics)

Expression profiles of candidate genes implicated in the migration and salinity acclimation in the threatened lampreys *Lampetra fluviatilis* and *L. planeri*

Catarina Mateus

Host researcher: Professor Walter Salzburger
Host Institute: Zoological Institute, University of Basel

Scientific report
December 2012

1. Purpose of the visit

The main objective of the proposed project work was to identify putative genes responsible for the marked shift in morphology and life history of the closely related species pair *Lampetra fluviatilis* and *L. planeri*, and thus contribute to the understanding of the relatedness between members of lamprey species pairs.
Also, it was my objective to learn and discuss several technical and practical aspects related to the application of next-generation sequencing technologies with the host researcher and his lab members, and share some ideas for future collaborations.

2. Description of the work carried out during the visit

As a result of the discussions with the host researcher and his lab members, we decided that the first step should be to perform restriction-site associated DNA sequencing (RAD sequencing) to look for differences in the genome of both species. After this, I then performed qPCR using RNA for three candidate genes putatively involved in the migratory activity.

2.1 RAD sequencing

DNA extraction was performed with the “DNeasy Blood & Tissue Kit” (Qiagen) following the manufacturer’s protocol. To obtain genetic markers I prepared a RAD tag library. Briefly, genomic DNA was diluted to a final concentration of 22.5 ng/µl and then digested with a restriction enzyme (Sbf1). The digest was then ligated to a P1 adapter for sample identification and the samples were then multiplexed to a single library. The pooled samples were randomly sheared with a Bioruptor sonicator, size-selected through gel extraction and then ligated to a second adapter (P2 adapter). The last step was a final RAD tag amplification, followed by purification and gel extraction.
The library combined RAD from 18 individuals of each species (36 individuals in total) and was single-end sequenced with 100 cycles in a separate lane on an Illumina HiSeq genome analyzer, following quality control. The resulting reads were filtered for quality and then sorted individually by barcode.

2.2 Gene expression

I extracted RNA from 10-12 samples of each species from three tissues, muscle, liver and brain. RNA was extracted using TRIZOL Reagent (Invitrogen). Tissues were homogenized with a BeadBeater (FastPrep-24; MP Biomedicals). DNase treatment following the DNA Free protocol (Ambion) was performed to remove any genomic DNA from the samples. The RNA concentration was measured with a Nanodrop 1000 spectrophotometer (Thermo Scientific) and subsequent reverse transcription was achieved by using the High Capacity RNA-to-cDNA kit (Applied Biosystems) following the manufacturer’s protocol.
The next step was to select a list of candidate genes identified as being implicated in the migratory activity. For the six selected genes, primers were developed using the

sea lamprey as reference and selecting conserved regions across species (using BLAST). I performed PCR reactions to optimize the primers for the target species and to select the tissue where each gene is expressed in higher levels, and I then resequenced the genes for the target species and developed specific qPCR primers for some of the selected genes. Subsequent Real-Time PCR analyses were conducted on a StepOnePlus Real-Time PCR System (Applied Biosystems) following the manufacturer's protocol. Final cDNA and primer concentrations of 1 ng/µl and 200 nM, respectively, were used together with the SYBR Green master (Rox) dye (Roche, Basel, Switzerland). Relative quantification of candidate gene expression was calculated using the comparative threshold (C_T) method, with the gene coding for ATP synthase as the housekeeping gene. For the reference and for each of the target genes I constructed a standard curve using dilutions of a known template and used this curve to determine the efficiency of the assay. An efficiency close to 100% (ideally between 90 and 105%) is the best indicator of a robust, reproducible assay. The coefficient of determination (R^2) should be above 0.98 and the replicates should give similar C_T values.

3. Description of the main results obtained

3.1 RAD sequencing

The Agilent 2100 Bioanalyzer was used for sizing, quantification and quality control of DNA after library preparation. As seen in figure 1 and table I, most of the fragments in the library are of the expected molecular weight and thus suitable for sequencing in the Illumina HiSeq genome analyzer.

Because the sea lamprey genome proved to be significantly different from our species (after BLAST 100 of our reads against the sea lamprey genome), we decided not to use the sea lamprey genome as a reference genome.

The software Stacks is being used for RAD sequencing analyses.

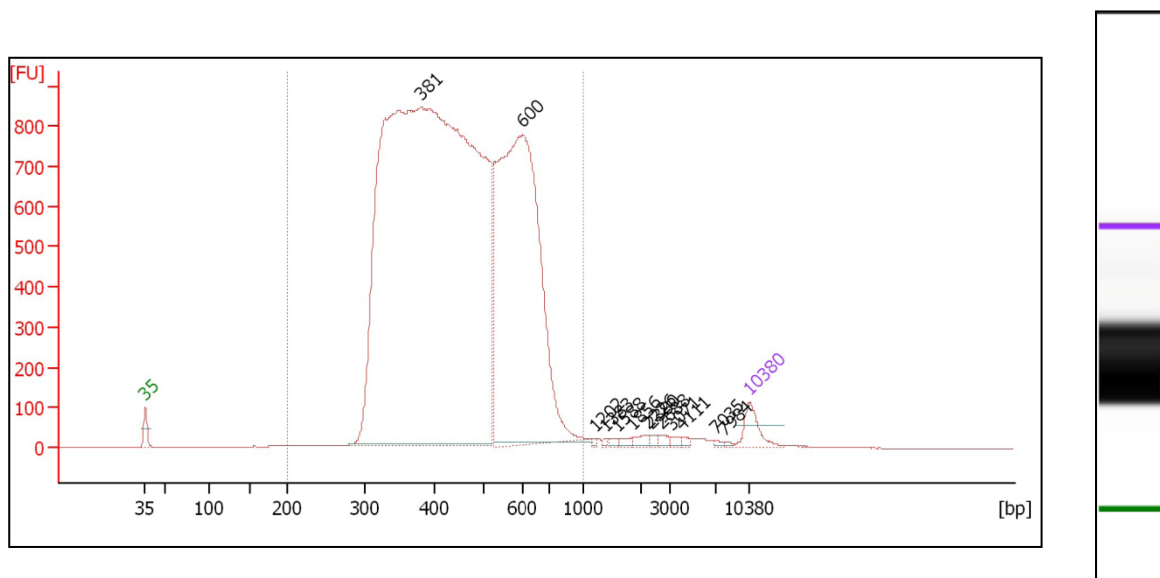

Figure 1 - Library electropherogram measured in an Agilent 2100 Bioanalyzer.

Table I – RAD library sizes and concentrations.

Peak		Size [bp]	Conc. [pg/µl]	Molarity [pmol/l]	Observations
1	▶	35	125.00	5'411.3	Lower Marker
2		381	8'053.38	31'984.6	
3		600	2'502.03	6'320.4	
4		1'202	6.03	7.6	
5		1'383	5.80	6.4	
6		1'588	9.70	9.3	
7		1'856	14.03	11.5	
8		2'226	19.97	13.6	
9		2'380	10.38	6.6	
10		2'688	13.91	7.8	
11		3'071	13.67	6.7	
12		4'111	9.14	3.4	
13		7'035	6.46	1.4	
14		7'984	4.93	0.9	
15	▶	10'380	75.00	10.9	Upper Marker

Note-number of peaks found: 13

3.2 Gene expression

Selected genes were those coding for endozepine (EDZP), cofilin 1 (CFL1), gonadotropin II beta subunit (GTH II), growth hormone (GH), lactate dehydrogenase (LTDH) and parvalbumin (PVBM). The genes showed different expression levels among tissues (e.g. fig. 2), and after optimization of the PCR conditions for the target species I selected GTH II, EDZP and PVBM for the qPCR experiments. Because RNA concentrations in the brain were low I used liver and muscle to the qPCR experiments.

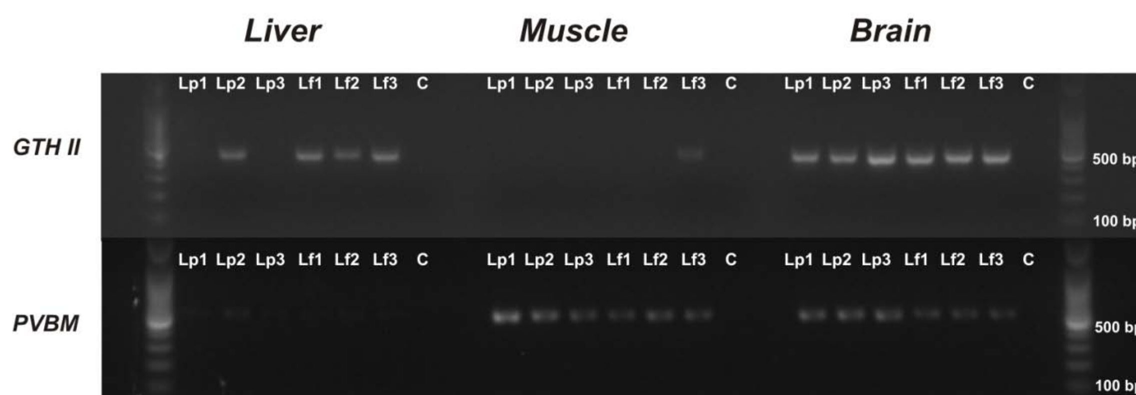

Figure 2 - Expression patterns of gonadotropin II beta subunit (GTH II) and parvalbumin (PVBM) for the liver, muscle and brain. For each PCR 15 ng of template cDNA was used and 3 µl PCR product was loaded on a 1.5% agarose gel. For each tissue, three samples of each *Lampetra planeri* (Lp) and *L. fluviatilis* (Lf) were amplified, and H_2O was used as negative control (C).

For the control, EDZP and PVBM the standard curves revealed high amplification efficiency and $R^2 > 0.98$, for both species. The gene coding for GTH II, however, had low efficiency (figs. 3 and 4).

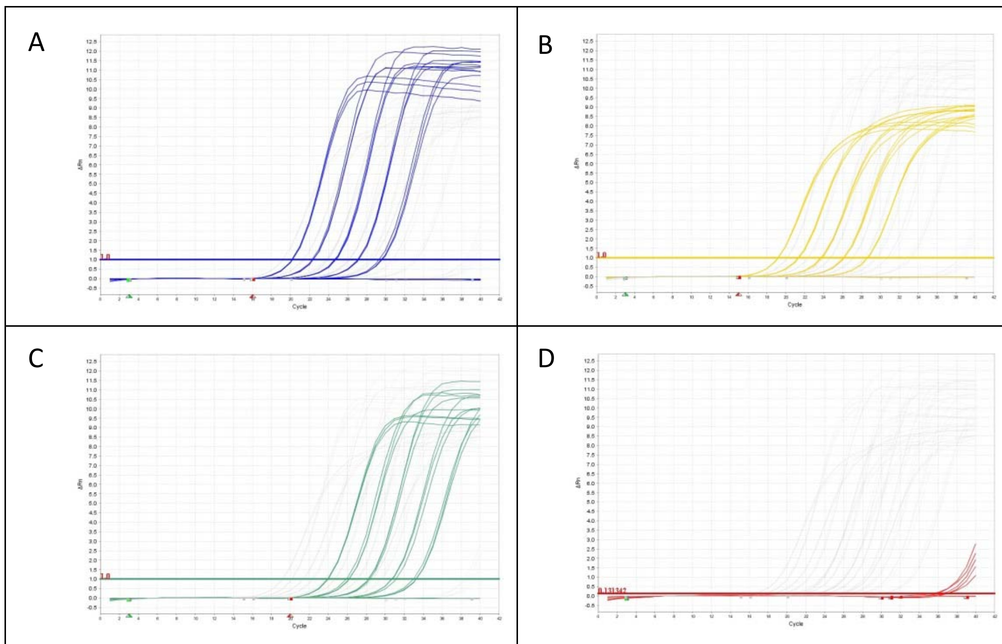

Figure 3 - Amplification plots for the four genes in *L. planeri.* A, ATP synthase; B, endozepine (EDZP); C, parvalbumin (PVBM); D, gonadotropin II beta subunit (GTH II).

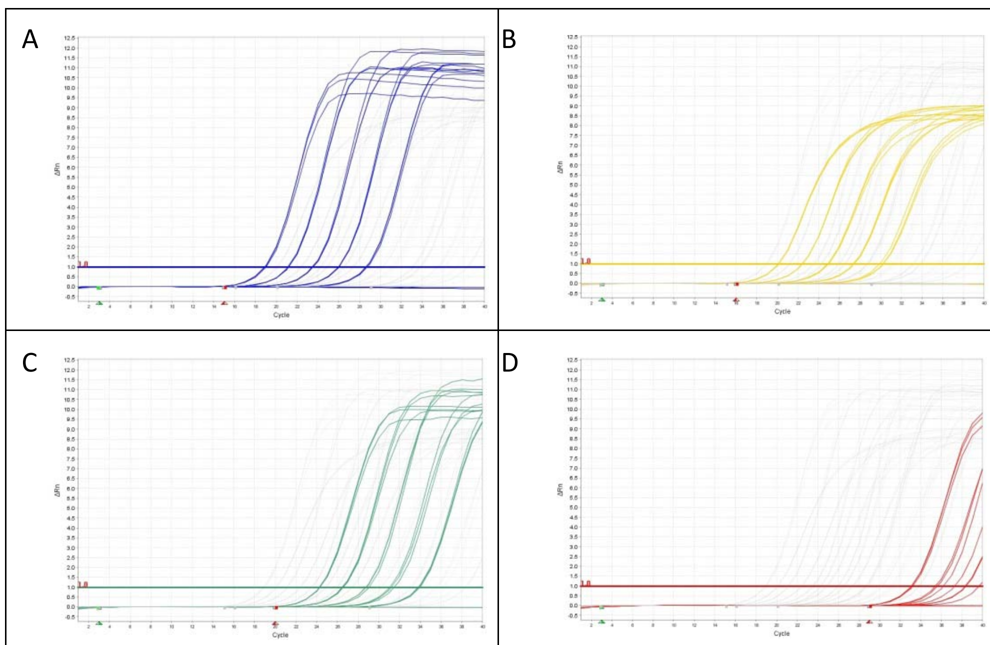

Figure 4 - Amplification plots for the four genes in *L. fluviatilis.* A, ATP synthase; B, endozepine (EDZP); C, parvalbumin (PVBM); D, gonadotropin II beta subunit (GTH II).

The gene expression levels and patterns of EDZP and PVBM are represented in Fig. 5. For both analyzed tissues (liver and muscle), PVBM showed the highest differences in expression between species, with higher expression in the resident species *L. planeri*.

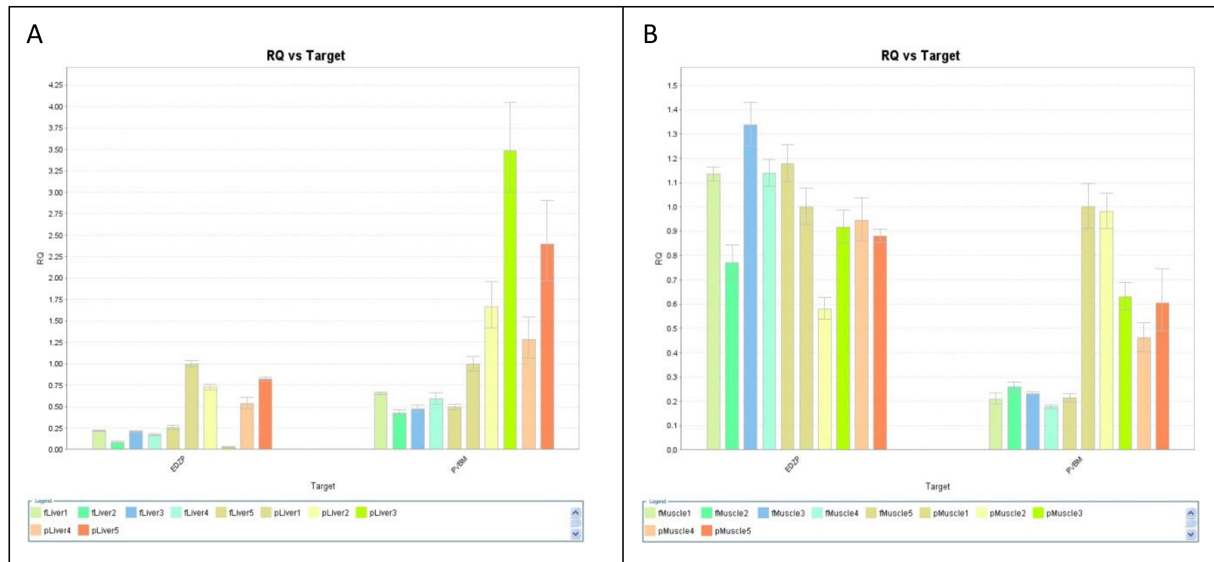

Figure 5 – Relative quantitation (RQ) plots representing the gene expression levels of endozepine (EDZP) and parvalbumin (PVBM) for liver (A) and muscle (B). For each gene five individuals each of *L. planeri* (p) and *L. fluviatilis* (f) were tested.

4. Future collaboration with host institution

In 2013 we intend submit a project to the Portuguese Foundation for Science and Technology entitled “TRANSCRIPT-Comparative transcriptomics of lampreys (genus *Lampetra*) with alternative life histories using next-generation sequencing techniques”, in collaboration with Professor Walter Salzburger (host researcher).

5. Projected publications to result from the grant

In the near future we intend to publish the results of the RAD sequencing attained in Basel and with the new project we will further explore the gene expression profiles in the species of interest.