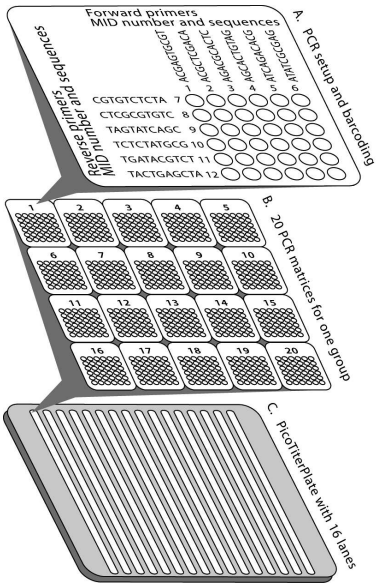
*Our research – your profit*

# Detection and mapping of mtDNA SNPs in Atlantic salmon using high throughput DNA sequencing

**Olafur Fridjonsson[1], Scott Tompsett[2], Kristinn Olafsson[1], Snaedis Bjornsdottir[1], Sonia Consuegra[2], David Knox[4], Carlos Garcia de Leaniz[3], Steinunn Magnusdottir[1], Gudbjorg Olafsdottir[1], Eric Verspoor[4], Ragnar Johannsson[1] and Sigridur Hjorleifsdottir[1]**

[1] Matís, Vínlandsleið 12, 113 Reykjavík, Iceland. [2] Institute of Biological Sciences, University of Wales, Aberystwyth, UK

[3] School of Biological Sciences, University of Wales Swansea, Swansea SA2 8PP, UK. [4] Freshwater Laboratory, Marine Scotland, Pitlochry, Scotland PH16 5LB, UK




## Implementation

546 salmon DNA samples (+30 charr and trout samples) were divided into 16 groups. Mitochondrial DNA fragments were amplified from each group in a matrix of 36 PCR reactions, using barcoded primers (MIDs) (A). A total of 20 matrices of PCR reactions, one for each mitochondrion fragment, were carried out for each group (B). Up to 720 amplicons from each group were pooled together in near equimolar concentrations. The 16 pools created, were used to generate single stranded DNA templates for the FLX sequencing. 16 amplified DNA bead -libraries, were loaded on PicoTiterPlate equipped with a sixteen-lane gasket, each library was assigned one lane (C). The combination of barcoded primers and a partitioned sequencing plate was used to assign each sequence read back to the original sample

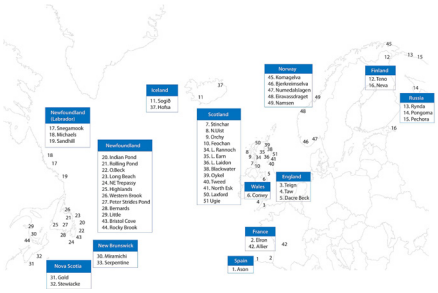
## SNP validation - Criteria

(i) >90% support from sequencing reads - a total of more than 10 supporting reads in both read directions - if less than 10 supporting reads, the SNP should be present in other samples with higher support

(ii) If <90% support from sequencing - total of more than 10 reads, in forward and reverse direction - the SNP present in other samples. SNPs only found in one sample rejected

## Objectives

- To design an optimal, cost-effective 454 pyro-sequencing protocol for screening large numbers of amplicons from a large number of individuals in the most efficient and cost-effective manner.
- To identify a set of minimally biased mtDNA SNPs for biodiversity and phylogeographic studies of the Atlantic salmon, which are representative, both of the genome and of species populations across the species trans Atlantic range.


- **20 mitochondrial amplicons** shown on the map (top) from **546 specimens** of Atlantic salmon, derived from across the species North Atlantic range (map below) sequenced in a single FLX Titanium run **(10920 PCR products)**
- Variants analyzed with the GS Amplicon Variant Analyzer Software

## Results

- Single PCR products of the expected sizes were obtained from the majority of the **11,520** reactions performed.
- The pyrosequencing yielded a total of **179,826,884** filter passed base pairs. An average of 11,546,081 bp was obtained for each of the 16 lanes on the picotiter plate, excluding one region only giving 6,635,668 bp.
- The majority of sequencing reads was about 400-420 bp corresponding to the size of the amplicons.
- Excluding the trout and charr sequences, a total of 10,765 sequences of 10920 were obtained or **98,6%**.
- The number of sequences supporting the SNPs per individual ranged from **4-60** sequences with the average of **28 ±12**.
- An alignment analysis yielded a total of 207 polymorphic loci, with 205 SNPs and 2 indels, thereof, **187** SNPs which have not been published.

### Target regions and number of SNPs

| Region | Size (bp) | Amplicon number | Amplicon Size | Number of SNP |
|---|---|---|---|---|
| DLOOPB | 422 | 1 | 422 | 21 |
| ND1 | 1160 | 2 | 422 | 13 |
| | | 3 | 406 | 5 |
| | | 4 | 363 | 8 |
| ND2 | 769 | 5 | 400 | 13 |
| | | 6 | 389 | 7 |
| COXI | 820 | 7 | 409 | 12 |
| | | 8 | 422 | 12 |
| COXII | 714 | 9 | 401 | 6 |
| | | 10 | 346 | 7 |
| ATP6 | 413 | 11 | 413 | 12 |
| ND3 | 402 | 12 | 402 | 10 |
| ND4 | 1180 | 13 | 400 | 10 |
| | | 14 | 401 | 12 |
| | | 15 | 414 | 12 |
| ND5 | 782 | 16 | 392 | 7 |
| | | 17 | 411 | 11 |
| CYTB | 1161 | 18 | 403 | 10 |
| | | 19 | 395 | 7 |
| | | 20 | 404 | 12 |
| Together | 7823 | | 8015 | 207 |

## Conclusions

The study generated a large mitochondrial SNP set that will be useful for individual genotyping and evolution studies of the Atlantic salmon.

Furthermore, the study shows that the approach described can be applied for high throughput sequencing of targeted regions in numerous individuals of other species of interest.