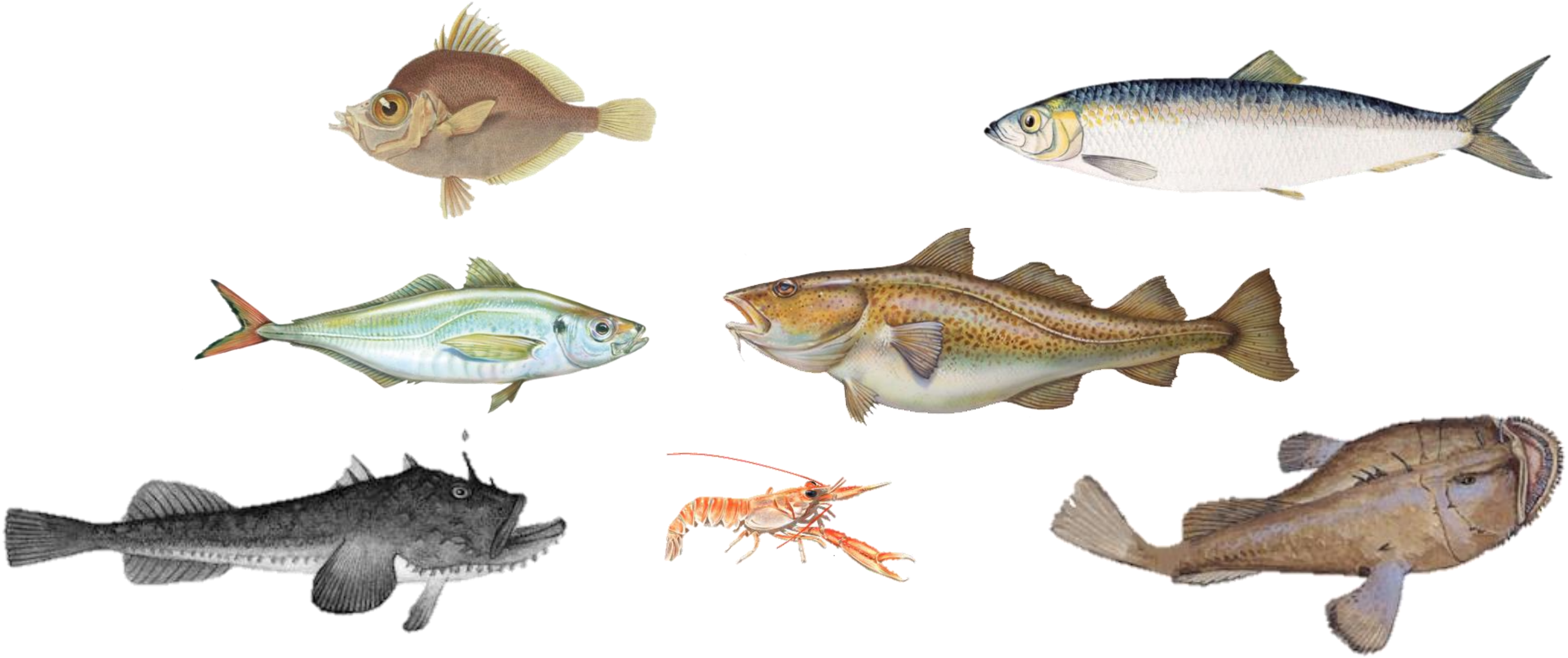
Fisheries Genetics for Stock Identification



NWWAC
1st March 2017, Paris

Edward Farrell & Jens Carlsson
University College Dublin, Ireland

Why is stock Identification essential?

Stock assessment assumptions

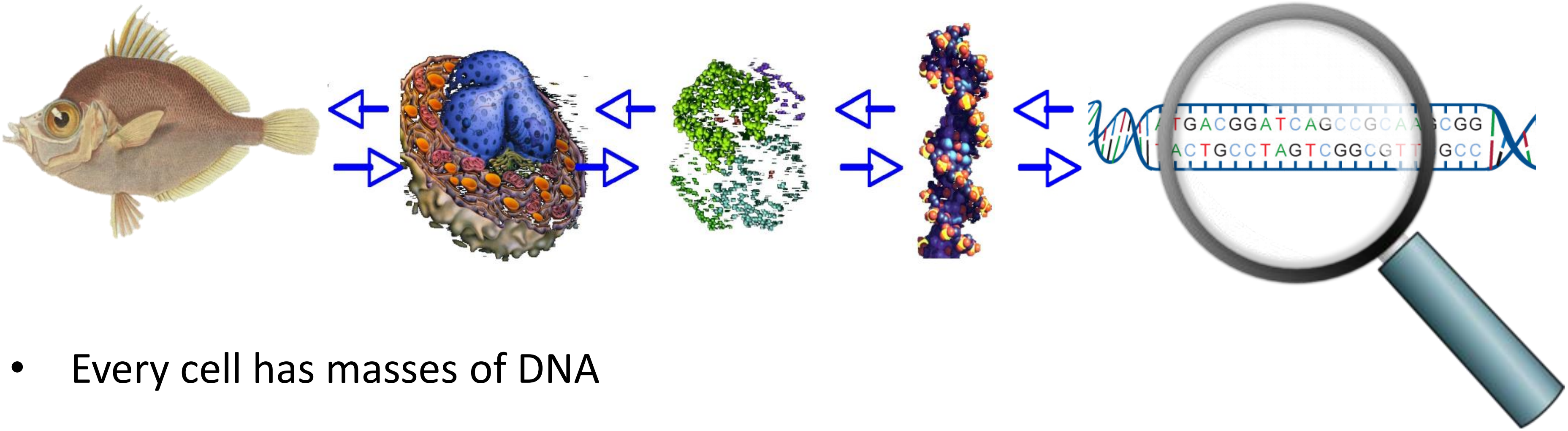
1. Stock is a closed population
    a) Gain through recruitment not immigration
    b) Loss through mortality not emigration
2. Data represents entire stock
    a) Even distribution of catches (fisheries dependent data)
    b) Survey covers entire stock area (fisheries independent data)

- Stock Identification is essential prerequisite for stock assessment
- Most assessments do not have good stock resolution

Stock Identification methods

- Indirect methods
  - Tagging = physical movement, migrations, mixing
  - Morphology and biology = physical appearance and characteristics
    - Combination of genetics and environment
- Direct methods
  - Genetics
    - Examining differences in genetic code
    - Widely used in medical and agricultural sciences, fisheries is lagging behind
    - Used to be slow, expensive and low power to detect stock structure
    - Recent advances have made it more applicable to stock Identification
    - **Next Generation Sequencing (NGS)**
    - **Genotyping by Sequencing (GBS)**

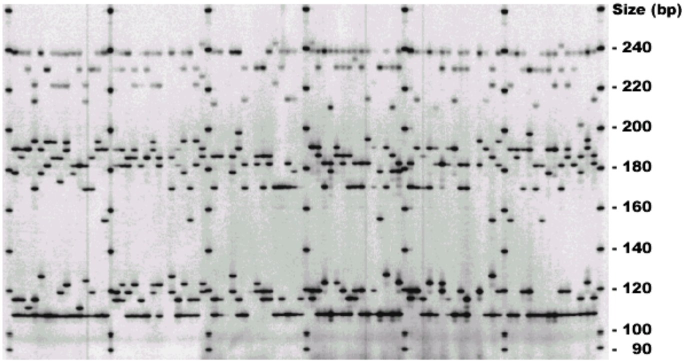
Basis of genetic stock Identification = DNA

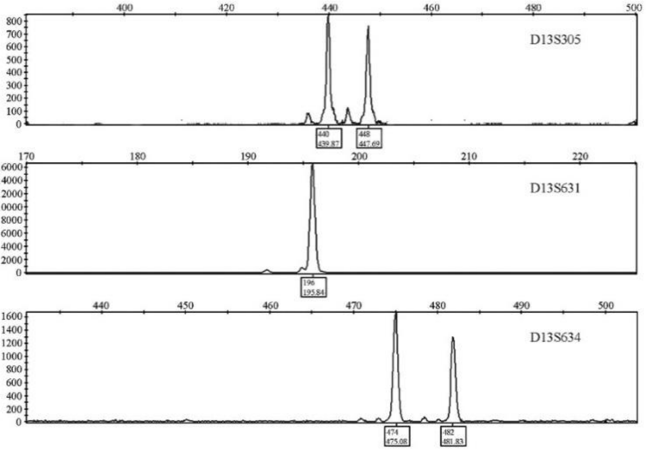


- Every cell has masses of DNA
- Blueprint of the organism = Simple 4 base code –ATGC
- 2 metres of DNA in every human cell
- Masses of potential info about individuals and populations
- Need tools to extract the relevant info
- Different studies require different info at different levels
- **Genetic Markers**

Traditional Fisheries Genetics

- Costly and difficult to develop markers
- Low numbers of informative markers
- Poor resolution
- Low reproducibility between labs
- One-off studies (MSc/PhD)
- Academic bias
- No annual monitoring
- Limited contribution to assessments

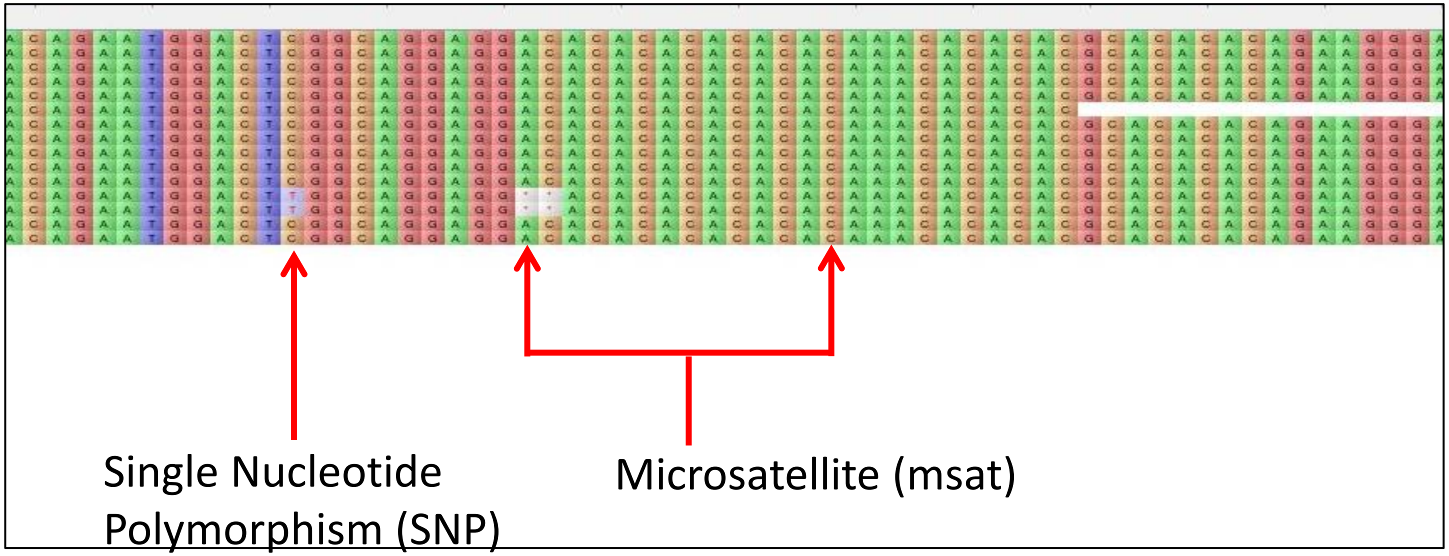




Next Generation Fisheries Genetics

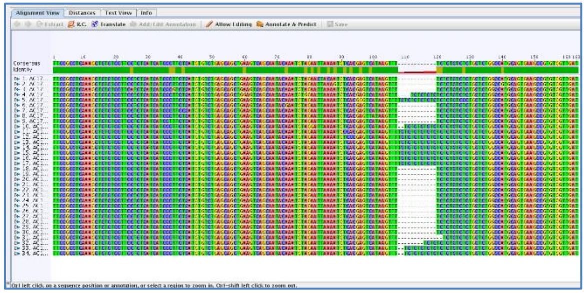
Current Development

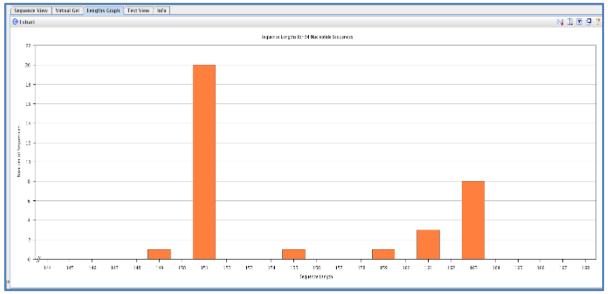
- **Cheaper, Faster, Better, More**
- Next Generation Sequencing (NGS)
- Genotyping By Sequencing (GBS)



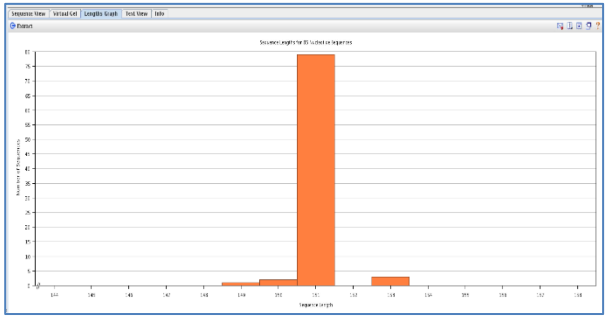
Single Nucleotide Polymorphism (SNP)

Microsatellite (msat)





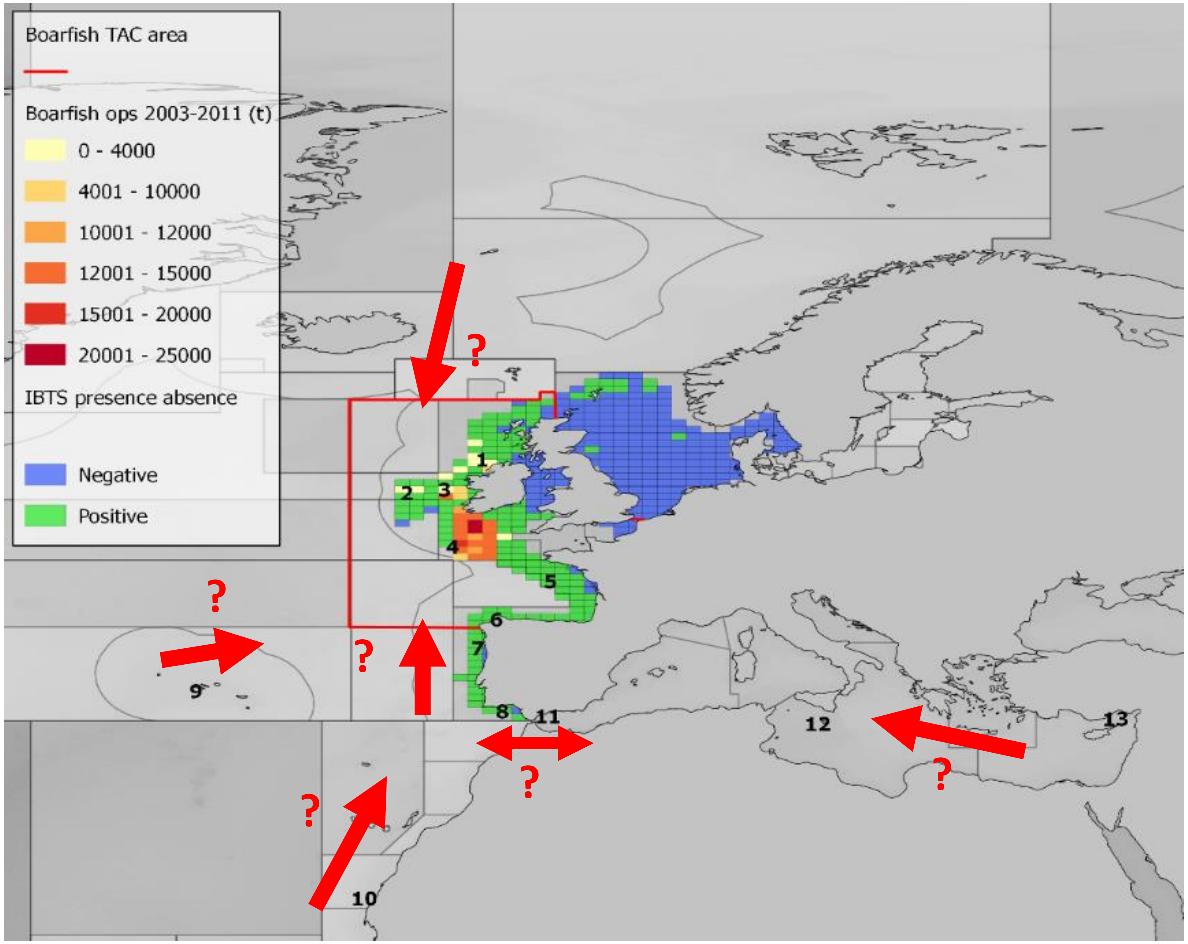
Heterozygote
151 163



Homozygote
151 151

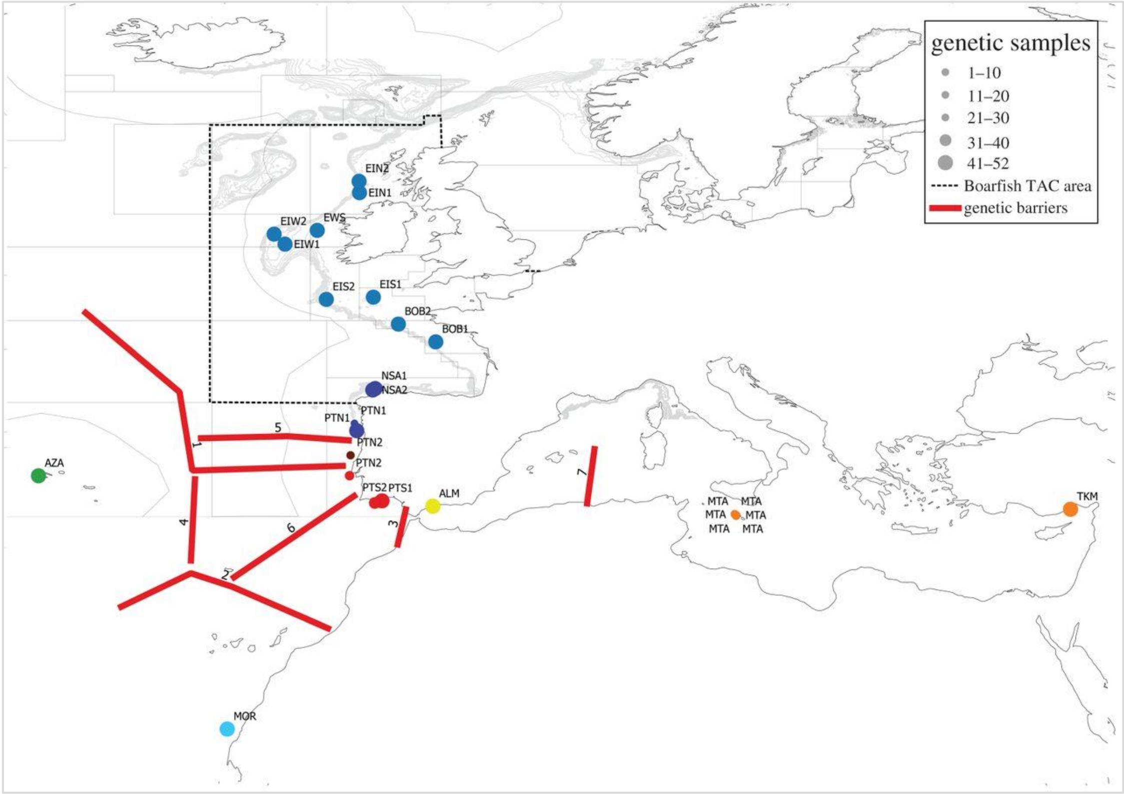
Example 1: Boarfish

- Does assessment/management area align with the population structure?
- Is there population structure within the assessment/management area?
- Is there immigration into the assessment/management area?



Example 1: Boarfish

- De novo generation of 85 microsatellite markers
- 960 samples screened in a single sequencing run
- Rapid identification of populations
- Fine scale resolution e.g. Portuguese mixing zone
- Results that can be used for better assessment and management
- Annual monitoring of the stock possible



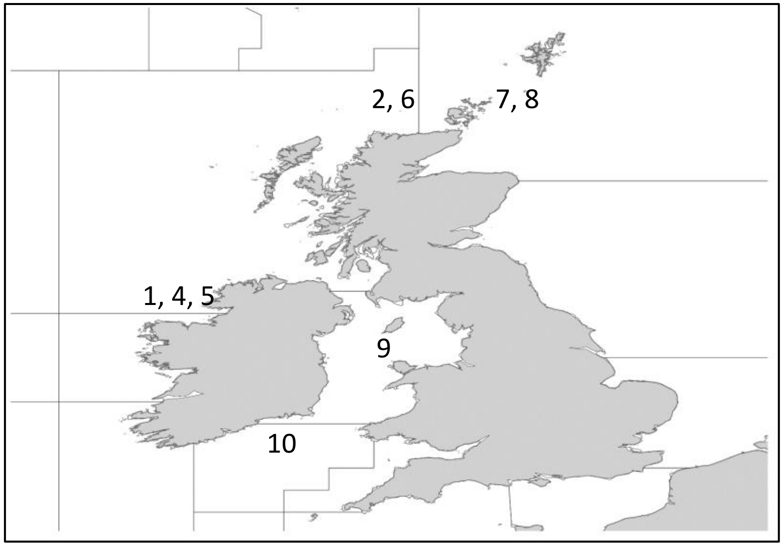
Example 1: Boarfish

- No previous genetic information available for boarfish
- 85 de novo microsatellites
- Not feasible using conventional methods
- 960 individuals @ 85 msats = 81,600 genotypes
- ABI platform (old method of capillary electrophoresis)
    - c. €4 per genotype
    - c. €326,400 in lab costs and at least 3 years
- Next Generation Genotyping-By-Sequencing
    - c. €15,000 in lab costs and 2 months C. €0.18 per genotype
- Further optimisation and automation is possible

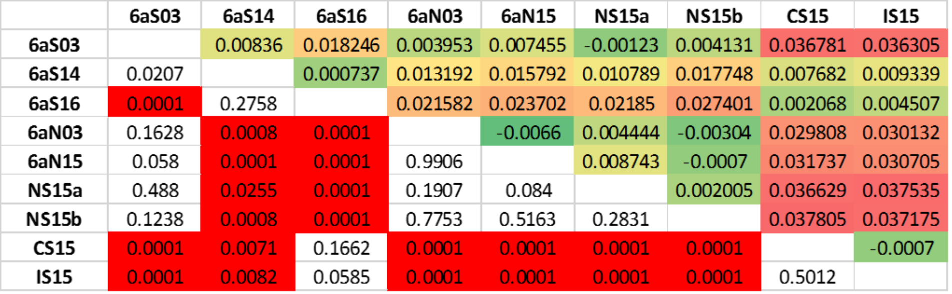
Example 2: 6a/7bc Herring

- 6aN versus 6aS & 7b,c
- Separate populations?
- Separate assessments?
- WESTHER (2003-2006) found no significant population structure in 6a
- KFO pilot study to reanalyse the WESTHER samples found significant structure
- Full scale project started 2016

#	Baseline samples	Code	Year	n
1	WESTHER Donegal	6AS03	2003	86
2	WESTHER Cape Wrath	6AN03	2003	84
3	WESTHER western Baltic	BLT	2004	22
4	VIaS	6AS14	2014	192
5	VIaS	6AS16	2016	192
6	VIaN	6AN15	2015	96
7	North Sea	NS15a	2015	96
8	North Sea	NS15b	2015	48
9	Irish Sea	IS15	2015	96
10	Celtic Sea	CS15	2015	96
	Total			**1008**

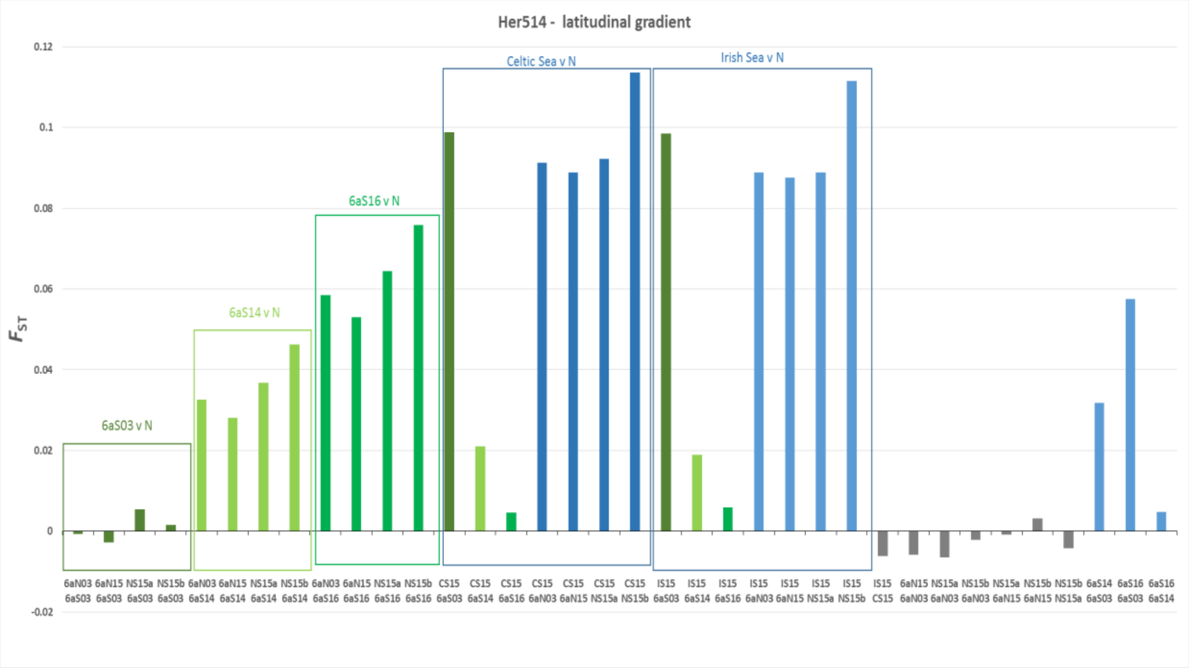


Example 2: 6a/7bc Herring



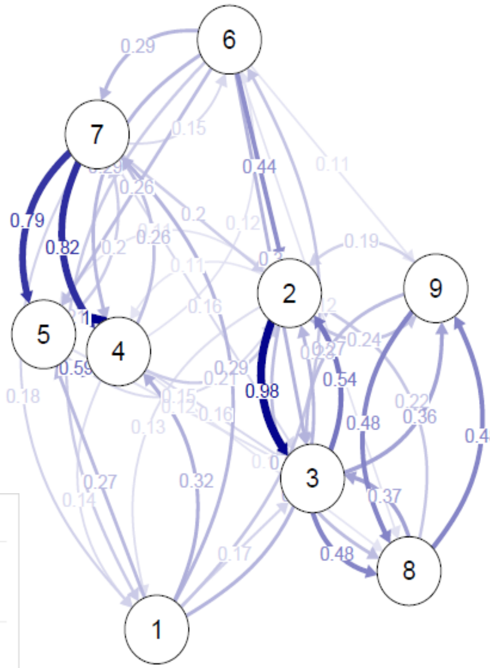
	6aS03	6aS14	6aS16	6aN03	6aN15	NS15a	NS15b	CS15	IS15
6aS03		0.00836	0.018246	0.003953	0.007455	-0.00123	0.004131	0.036781	0.036305
6aS14	0.0207		0.000737	0.013192	0.015792	0.010789	0.017748	0.007682	0.009339
6aS16	0.0001	0.2758		0.021582	0.023702	0.02185	0.027401	0.002068	0.004507
6aN03	0.1628	0.0008	0.0001		-0.0066	0.004444	-0.00304	0.029808	0.030132
6aN15	0.058	0.0001	0.0001	0.9906		0.008743	-0.0007	0.031737	0.030705
NS15a	0.488	0.0255	0.0001	0.1907	0.084		0.002005	0.036629	0.037535
NS15b	0.1238	0.0008	0.0001	0.7753	0.5163	0.2831		0.037805	0.037175
CS15	0.0001	0.0071	0.1662	0.0001	0.0001	0.0001	0.0001		-0.0007
IS15	0.0001	0.0082	0.0585	0.0001	0.0001	0.0001	0.0001	0.5012	

- Analysis ongoing
- Informative markers identified



Her514 - latitudinal gradient

Relative migration network
(Filter threshold = 0.1; Gst)



1: 6aS03,
2: 6aS14,
3: 6aS16,
4: 6aN03,
5: 6aN15,
6: NS15a,
7: NS15b,
8: CS15,
9: IS15,

Where are the informative markers?

Any additional markers near them?

Cherry-pick markers from other studies

- Draft Herring genome (Martinez Barrio *et al*., 2016)
- Identify relevant areas
- 70 SNPs related to spawning time
- Now screened a subset of baseline samples
- Very preliminary results

eLIFE
elifesciences.org

RESEARCH ARTICLE

The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing

Alvaro Martinez Barrio[1,2†], Sangeet Lamichhaney[1†], Guangyi Fan[3,4†], Nima Rafati[1†], Mats Pettersson[1], He Zhang[4,5], Jacques Dainat[1,6], Diana Ekman[7], Marc Höppner[1,6], Patric Jern[1], Marcel Martin[7], Björn Nystedt[2], Xin Liu[4], Wenbin Chen[4], Xinming Liang[4], Chengcheng Shi[4], Yuanyuan Fu[4,8], Kailong Ma[4], Xiao Zhan[4], Chungang Feng[1], Ulla Gustafson[9], Carl-Johan Rubin[1], Markus Sällman Almén[1], Martina Blass[10], Michele Casini[11], Arild Folkvord[12,13,14], Linda Laikre[15], Nils Ryman[15], Simon Ming-Yuen Lee[3], Xun Xu[4], Leif Andersson[1,9,16*]

[1]Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden; [2]Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden; [3]State Key Laboratory of Quality Research in Chinese Medicine, Institute of Chinese Medical Sciences, University of Macau, Macau, China; [4]BGI-Shenzhen, Shenzen, China; [5]College of Physics, Qingdao University, Qingdao, China; [6]Bioinformatics Infrastructure for Life Sciences, Uppsala University, Uppsala, Sweden; [7]Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden; [8]School of Biological Science and Medical Engineering, Southeast University, Nanjing, China; [9]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden; [10]Department of Aquatic Resources, Institute of Coastal Research, Swedish University of Agricultural Sciences, Öregrund, Sweden; [11]Department of Aquatic Resources, Institute of Marine Research, Swedish University of Agricultural Sciences, Lysekil, Sweden; [12]Department of Biology, University of Bergen, Bergen, Norway; [13]Hjort Center of Marine Ecosystem Dynamics, Bergen, Norway; [14]Institute of Marine Research, Bergen, Norway; [15]Department of Zoology, Stockholm University, Stockholm, Sweden; [16]Department of Veterinary Integrative Biosciences, Texas A&M University, Texas, United States

***For correspondence:** leif.andersson@imbim.uu.se

†These authors contributed equally to this work

Competing interests: The authors declare that no competing interests exist.

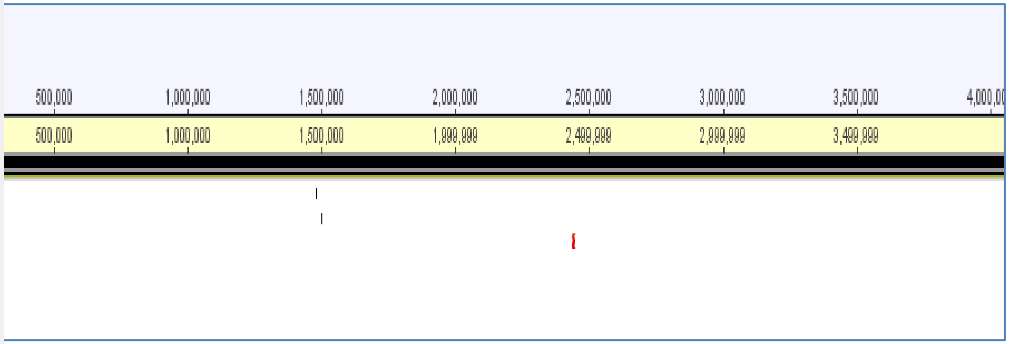
Funding: See page 25

Received: 05 October 2015
Accepted: 06 April 2016
Published: 03 May 2016

Reviewing editor: Magnus Nordborg, Vienna Biocenter, Austria

Ⓒ Copyright Martinez Barrio et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract Ecological adaptation is of major relevance to speciation and sustainable population management, but the underlying genetic factors are typically hard to study in natural populations due to genetic differentiation caused by natural selection being confounded with genetic drift in subdivided populations. Here, we use whole genome population sequencing of Atlantic and Baltic herring to reveal the underlying genetic architecture at an unprecedented detailed resolution for both adaptation to a new niche environment and timing of reproduction. We identify almost 500 independent loci associated with a recent niche expansion from marine (Atlantic Ocean) to brackish waters (Baltic Sea), and more than 100 independent loci showing genetic differentiation between spring- and autumn-spawning populations irrespective of geographic origin. Our results show that

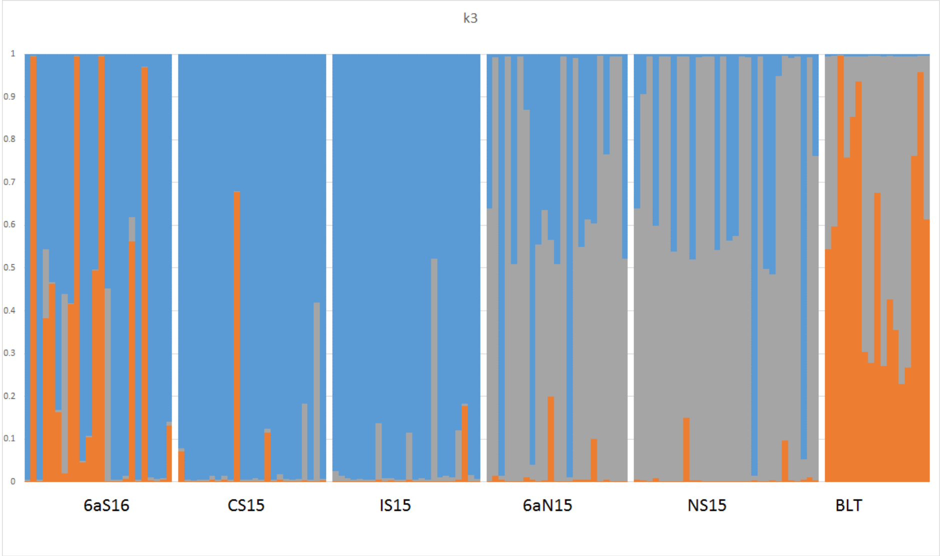


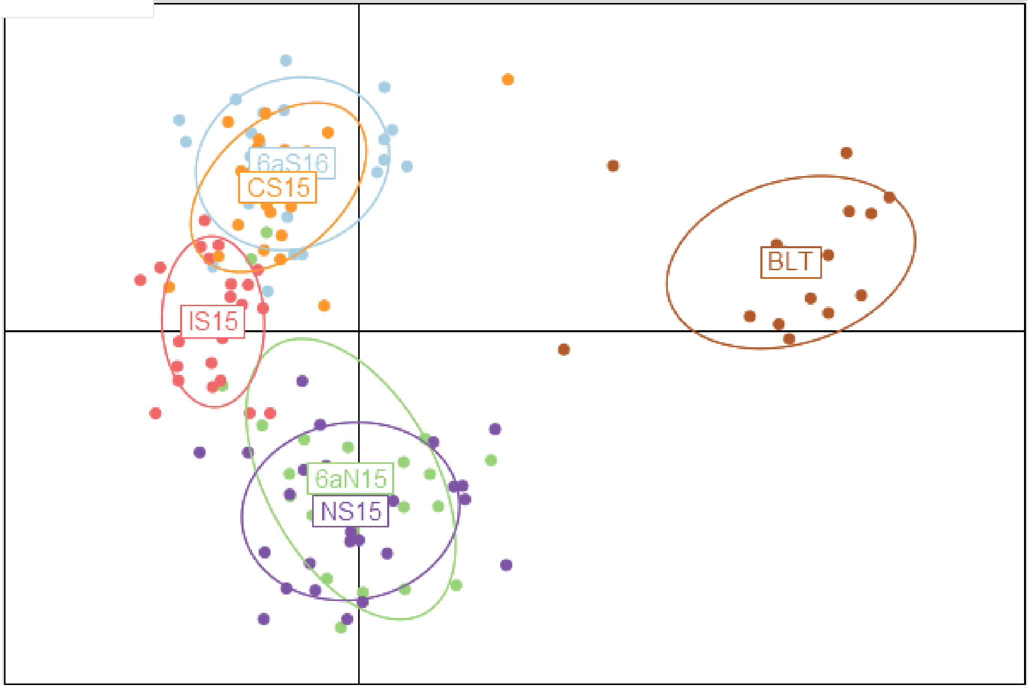
Very preliminary herring SNP results

- 70 SNPs related to spawning time
- 6aS16 (24), 6aN15 (23), IS15 (24), CS15 (24), NS15 (30), BLT (17)
- 56/70 SNPs successfully scored

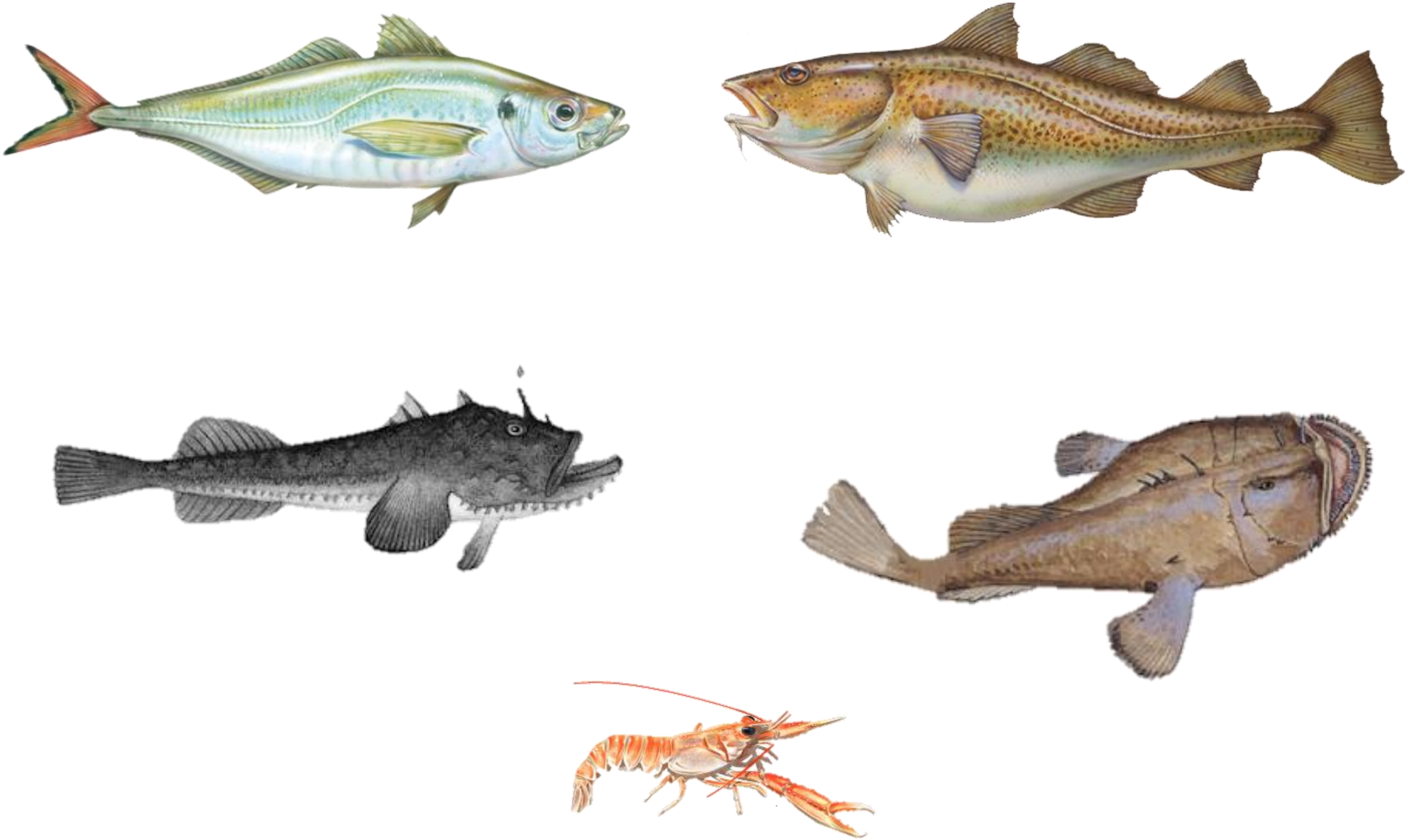
	6aS16	6aN16	IS15	CS15	NS15	BLT
6aS16		0.214343	0.136576	0.108908	0.289566	0.358375
6aN16	0.0001		0.281802	0.328253	0.003487	0.433712
IS15	0.0001	0.0001		0.12493	0.370546	0.634549
CS15	0.0019	0.0001	0.0001		0.419047	0.615281
NS15	0.0001	0.2741	0.0001	0.0001		0.465325
BLT	0.0001	0.0001	0.0001	0.0001	0.0001	

Very preliminary herring SNP results





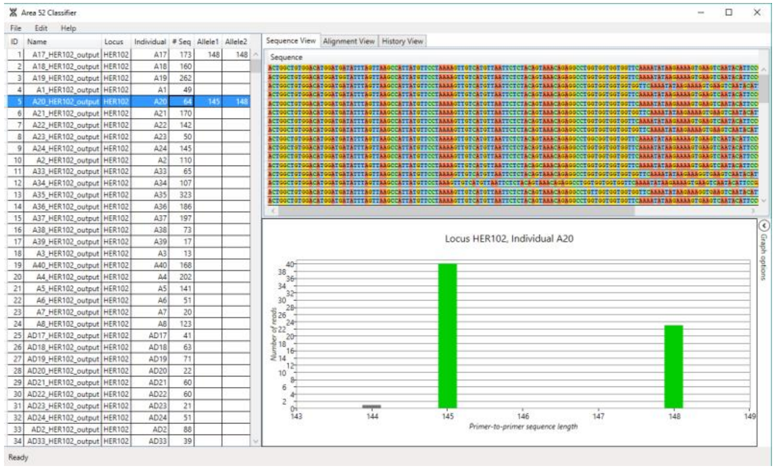
Other stock ID Projects



Further Optimisation

Upscale

- Streamlined sample collection and processing
- Automated DNA extraction
- High-throughput lab protocols
- Automated genotyping



ID	Name	Locus	Individual	# Seq	Allele1	Allele2
1	A17_HER102_output	HER102	A17	173	148	148
2	A18_HER102_output	HER102	A18	160		
3	A19_HER102_output	HER102	A19	262		
4	A1_HER102_output	HER102	A1	49		
5	A20_HER102_output	HER102	A20	64	145	148
6	A21_HER102_output	HER102	A21	170		
7	A22_HER102_output	HER102	A22	142		
8	A23_HER102_output	HER102	A23	50		
9	A24_HER102_output	HER102	A24	145		
10	A2_HER102_output	HER102	A2	110		
11	A33_HER102_output	HER102	A33	65		
12	A34_HER102_output	HER102	A34	107		
13	A35_HER102_output	HER102	A35	323		
14	A36_HER102_output	HER102	A36	186		
15	A37_HER102_output	HER102	A37	197		
16	A38_HER102_output	HER102	A38	73		
17	A39_HER102_output	HER102	A39	17		
18	A3_HER102_output	HER102	A3	13		
19	A40_HER102_output	HER102	A40	168		
20	A4_HER102_output	HER102	A4	202		
21	A5_HER102_output	HER102	A5	141		
22	A6_HER102_output	HER102	A6	51		
23	A7_HER102_output	HER102	A7	20		
24	A8_HER102_output	HER102	A8	123		
25	AD17_HER102_output	HER102	AD17	41		
26	AD18_HER102_output	HER102	AD18	63		
27	AD19_HER102_output	HER102	AD19	71		
28	AD20_HER102_output	HER102	AD20	22		
29	AD21_HER102_output	HER102	AD21	60		
30	AD22_HER102_output	HER102	AD22	60		
31	AD23_HER102_output	HER102	AD23	21		
32	AD24_HER102_output	HER102	AD24	51		
33	AD2_HER102_output	HER102	AD2	88		
34	AD33_HER102_output	HER102	AD33	39		

Develop

- Sampling individual fish is slow
- Why not sample all fish at once?
- Take water sample from RSW tank
- eDNA and ‘real-time’ stock ID

The Future?

- Reliable Stock Identification
- Robust and updated baselines
- Standardised and regular sampling
- Monitoring of stock mixing
- Current and relevant data for assessments
- Fast turnaround (months not years)

How to achieve this?

- Strong collaboration
  - Industry
  - Institutes
  - Tell us what you need
- Sustained funding
- Staying up to date with new technology

Real time monitoring