Electronic appendix for paper

Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring

Daniel E. Ruzzante, Stefano Mariani, Dorte Bekkevold, Carl Andre, Henrik Mosegaard, Lotte A. W. Clausen, Thomas G. Dahlgren, William F. Hutchinson, Emma M. C. Hatfield, Else Torstensen, Jennifer Brigham, E. John Simmonds, Linda Laikre, Lena C. Larsson, René J. M. Stet, Nils Ryman, Gary R. Carvalho

Molecular analysis: The genetic work was conducted in three laboratories: The microsatellite DNA work for spawning samples from locations 1-8 and summer mixed feeding aggregation samples from the northern North Sea (Shetland) was conducted at the Molecular Ecology and Fisheries Genetics Laboratory of the University of Hull (UK); the microsatellite work for spawning samples from locations 9-13 was conducted at the Tjärnö Marine Biological Laboratory of Göteborg University (Sweden), and the microsatellite work for spawning samples from locations 14-18 as well as for mixed feeding aggregation summer and winter samples from Skagerrak was conducted at the Department of Inland Fisheries of the Danish Institute for Fisheries Research (Denmark). A total of nine microsatellite loci were screened (Table A3) and details of extraction, amplification and scoring methods are available in Mariani et al (2005) and Bekkevold et al (2005).

Our results indicate there are three major genetic groups of herring which correspond almost exactly to the clusters of samples analysed by the three laboratories. Locations 1-8 were all analysed in England (Univ of Hull) and all fall in cluster 1, locations 14-18 were all analyzed in Denmark (DIFRES) and these represent the entire cluster 3; finally, some of the samples analyzed in Sweden (locations 9 and 10) fall into cluster 1, while some (locations 11-13) represent the entire cluster 2. Confidence in these results is based on the following points:

Firstly, as stated above, all mixed aggregation samples from the Skagerrak (summer and winter samples from 2002 and 2003) were analysed in Denmark. Despite this, the various MSA computations consistently assigned a conspicuous proportion of individuals from the mixed samples to the Skagerrak locations analysed in Sweden (locations 11-13) and to the North Sea Autumn spawners analysed in England. Were the distinct genetic composition of the three groups of samples a result of laboratory artifact, it is extremely unlikely any of the mixed aggregation samples (analysed in Denmark) would have been assigned to the group of spawning locations analysed in Sweden or England. Furthermore, according to the otolith readings, all individuals in the mixed aggregations that were assigned to the Skagerrak locations 11-13 were “spring-hatched” herring as would be expected for herring from this region. Had the assignments been the result of artifacts it is to be expected, just by chance, that some of the Skagerrak-assigned individuals would have been autumn-hatched, yet this is not the case. Similarly, individuals in the Skagerrak summer and winter mixed aggregations that were assigned to the North Sea were mostly autumn-hatched herring as would have been expected given

what is known of the biology of herring in the region. Had the MSA outcome been the result of laboratory artifacts, one would expect little, if any, correspondence between region of potential origin and hatching season.

Secondly, in the multidimensional scaling analysis, the samples from locations 11-13 (Skagerrak) cluster in the middle, between the group of samples from the North Sea (cluster 1) and the cluster of samples from the western Baltic (cluster 3). Should the distinctiveness among the three clusters of spawning samples be the artifactual result of inter-laboratory differences, one would not necessarily expect such a clustering. The fact that herring from the Skagerrak/Kattegat region cluster in the middle, is consistent with the hypothesis that they are genetically intermediate between those in the North Sea and those in the western Baltic. We stress that a similar situation has been described for other species in the region, including Atlantic cod (*Gadus morhua*, Nielsen et al 2003) MEC 12: 1497-1508) and turbot (*Scophthalmus maximus*, Nielsen et al 2004 MEC 13: 585-595) for each of which a hybrid zone has been postulated.

Thirdly (as noted above) the correspondence between clusters and laboratories is not complete as there are two locations analysed in Sweden that group with cluster 1, comprising all the locations analysed in England.

We point out that we designed our analyses specifically to ensure consistency among analytical platforms and have spent considerable effort cross-calibrating our genotyping results. Prior to the study, we blindly genotyped in each of the three laboratories, ten

individuals from each of four previously collected samples (total of 40 individuals). After initial calibration of allele sizes there were no genotyping inconsistencies in these 40 X 9 (individuals per loci) triple scorings (720 alleles observed across all 9 loci). These 40 individuals were not included in the current study, as none was ripe-and-running, an important condition to ensure the sampling of spawning individuals collected on their spawning grounds. Based on these initial genotyping results, we selected a series of test individuals for each locus that possessed alleles spanning the size range observed among the 40 individuals. These test samples were run alongside all samples in all three laboratories and sized using a standard, commercially available size ladder.

To further ensure consistency, later in the project we triple analyzed 6 other individuals chosen from a location outside the region (NW Atlantic). In this analysis, 1 allele out of a total of 108 observed across all 9 loci differed between English and Swedish/Danish laboratories (latter 2 laboratories were identical). Taken together the two calibration exercises provide an error rate of 1/ (720+108) = 0.0012. Such a low level of scoring error coupled with the absence of any trend across loci and/or samples indicates that the observed levels of population differentiation are not due to genotyping errors or bias. Despite our calibration and quality control efforts, it is possible that the ability to distinguish alleles of intermediate size (i.e. alleles differing by 2 bp from neighboring alleles in otherwise tetranucleotide loci) may have differed across platforms. There were a total of 89 such alleles scored over all 9 loci, and they appeared in 249 individuals out of the nearly 6000 individuals examined in the spawning and mixed aggregation samples. Collectively thus, these intermediate-size alleles represent < 0.0024 of all alleles scored

over loci and individuals. Overall and pairwise F_{ST} estimates among the three broad regions (North Sea, Kattegat, and Western Baltic) estimated before and after binning (rounding up) these alleles of intermediate size varied only at the fifth decimal place (e.g. overall F_{ST} estimates before and after binning are, respectively, F_{ST}= 0.00936 and F_{ST}= 0.00941). Any potential inconsistency across platforms in the scoring of these intermediate-size alleles is therefore not expected to affect the estimates of contributions to the mixed aggregations and their confidence intervals, and thus, our conclusions. In addition to the above, the software Microchecker (Van Oosterhout *et al.* 2004) was routinely employed by all three laboratories to test for heterozygote deficiencies arising from inconsistent genotyping of stutters and the presence of large allele dropout and null alleles.

Otolith age and hatching season error rate estimation: Age estimation from otolith winter rings is subject to error. In an effort to estimate this error for North Sea herring, an otolith exchange program took place among relevant institutions. Age readers in Scotland and Denmark, the main age readers in our study, achieved 94% consistency in age estimation for North Sea herring. Overall consistency across 8 readers involved in this exchange program (including readers not involved in our study) was 89% (ICES 2002).

The determination of spawning season from otolith microstructure is also subject to error and we have examined this aspect of our work (Clausen, Bekkevold, Hatfield & Mosegaard unpublished data). Hatching season determination was conducted at the Institute of Marine Research (IMR) in Norway and at the Danish Institute for Fisheries

Research in Denmark (DIFRES). Cross reader comparisons of their hatch season assignments showed a correspondence of 97% for spring spawning herring, 93% for autumn spawning herring and 68% for winter spawning individuals (Clausen, Bekkevold, Hatfield & Mosegaard unpublished data). The overlapping seasons of the autumn and winter spawning herring (Zijlstra, 1969; Burd and Howlett, 1974) and the gradual change in the otolith microstructure from autumn/late-autumn hatch type to winter hatch-type may cause mis-classification of early hatched winter spawners as autumn-hatch-types by visual inspection (Clausen, Bekkevold, Hatfield & Mosegaard unpublished data). The opposite mis-classification (autumn hatched misclassified as winter-hatched) is less common as the abundant autumn-hatch-type is more frequently seen in routine laboratory work and is therefore easily recognizable.

Appendix Table A1. Baseline ('learning') samples used in the mixed stock analysis. Locality numbers refer to Figure 1. Region numbers refer to groupings used in the mixed-stock analysis: Region 1 (North Sea autumn spawners, English Channel winter spawners and Norwegian spring spawners). Region 2 (Skagerrak spring spawners). Region 3 [Spring spawners from the Kattegat and western Baltic (Kattegat, Inner Danish Waters and Rügen)]. Samples based on combined temporal samples are indicated by two sampling dates. Details about DNA extraction and genotyping procedures are found in Mariani et al (2005) for samples collected at localities 1-10 , and in Bekkevold et al (2005) for samples collected at localities 9-18.

Locality name	Locality no.	Region no.	Latitude/ Longitude	Sampling date	Sample size	Mean age	ICES Region	Spawning season

Cape Wrath	1	1	58°38'N / 5°13'W	21/08/02	84	3.75	VIa	Autumn
				23/08/03	96	3.38		
Whiten Head	2	1	58°36'N / 4°20'W	01/09/02	99	3.00	VIa	Autumn
				01/09/03	86	3.12		
Shetland	3	1	60°29'N / 1°40'W	18/08/02	90	3.91	IVa	Autumn
Orkney	4	1	59°12'N / 1°40'W	12/08/02	106	2.71	IVa	Autumn
				21/08/03	85	3.58		
Aberdeen	5	1	57°42'N / 1°27'W	22/08/03	91	3.92	IVb	Autumn
Berwick	6	1	56°18'N / 0°58'W	20/08/02	100	2.72	IVb	Autumn
Flamborough	7	1	54°34'N / 0°27'W	14/09/02	97	3.52	IVb	Autumn
				17/09/03	77	3.37		
English Channel	8	1	50°7'N / 0°25'E	23/11/01	80	4.06	VIId	Winter
				23/11/03	63	3.93		
Møre	9	1	62°78'N / 06°08'E	17/02/03	78	8.31	IIa	Spring
Karmøy	10	1	59°14'N / 05°10'E	14/03/02	100	7.96	IVa	Spring
Tjöme	11	2	59°35'N / 10°55'E	13/03/02	120	4.30	IIIa	Spring
				04/03/03	116	5.22		
Måseskär	12	2	58°32'N / 11°32'E	19/03/02	100	3.82	IIIa	Spring
				14/03/03	99	4.23		
Flatbrotten	13	2	58°32'N / 11°25'E	07/03/02	100	3.84	IIIa	Spring
				19/03/03	100	3.99		
Limfjord	14	3	57°06'N / 10°06'E	22/05/03	99	3.11	IIIa	Spring
Kattegat	15	3	55°73'N / 11°37'E	06/05/02	44	3.91	IIIc/22	Spring
				030/4/03	99	4.40		
Kolding Fjord	16	3	55°49'N / 09°54'E	12/04/02	100	4.10	IIIc/22	Spring
				05/04/03	70	4.10		
Lillebælt	17	3	55°45'N / 09°72'E	07/04/03	100	5.95	IIIc/22	Spring
Rügen	18	3	54°23'N / 13°43'E	22/03/02	100	5.69	IIIc/24	Spring
				18/04/02	98	5.67		
				06/05/02	100	4.55		
				24/04/03	100	5.72		
				06/05/03	100	5.12		

Table A2. Samples of herring mixed feeding and overwintering aggregations (mixture samples). Distributions of spawning types (numbers of fish spawned in spring (S), autumn (A) and winter (W), and cases where type could not be determined (U) are given per sample. Although 492 and 472 herring from the northern North Sea mixed feeding aggregations sampled in 2002 and 2003, respectively, were sampled and examined for season hatched, the actual numbers used in the mixed stock analysis are, respectively, N=414 and N=310 (See Fig. 1 and Table A4).

Region	Year /season	Location	Latitude/Longitude	Sampling date	N	Individuals spawned in:			
						S	A	W	U
Northern	**2002**		60º03' N / 0º04' E	06/07/2002	100	10	46	33	11
North Sea	**(summer)**		60°21' N / 0°32' W	06/07/2002	100	2	75	16	7
(Shetland)			60°25' N / 0°37' E	07/07/2002	100	5	79	9	7
			60°31' N / 0°52' W	07/07/2002	100	2	85	10	3
			60°33' N / 0°33' W	07/07/2002	92	0	85	2	5
	Total N				**492**	**19**	**370**	**70**	**33**
	2003		60°03' N / 0°50' W	07/07/2003	100	2	66	29	3
	(summer)		60°24' N / 0°45' E	08/07/2003	100	4	77	12	7
			60°18' N / 0°06' E	08/07/2003	100	1	87	4	8
			60°33' N / 0°07' E	09/07/2003	100	2	71	23	4
			60°41' N / 0°42' W	10/07/2003	72	0	70	2	0
	Total N				**472**	**9**	**371**	**70**	**22**
Skagerrak	**2002**	Eastern North Sea	57°53' N / 6°35' E	29/06/02	200	92	29	40	39
(Summer)		Western Skagerrak	58°2' N / 8°22' E	01/07/02	200	34	42	21	103
		Central Skagerrak	58°37' N / 9°41' E	03/07/02	100	26	3	0	71
		Southern Skagerrak	58°10' N / 10°7' E	04/07/02	100	39	5	2	54
	Total N				**600**	**191**	**79**	**63**	**267**
	2003	Eastern North Sea	57°41' N / 6°55' E	02/07/03	100	42	49	9	0
		Western Skagerrak	58°13' N / 8°57' E	04/07/03	100	38	57	5	0

		Central Skagerrak	58°34' N / 9°57' E	06/07/03	100	73	22	3	2
		Southern Skagerrak	58°10' N / 10°5' E	07/07/03	100	38	57	5	0
	Total N				**400**	**191**	**185**	**22**	**2**
Skagerrak	**2002**	Grimstad	58°20' N / 8°33' E	17/12/02	100	100	0	0	0
(Winter)		Risør	58°44' N / 9°15' E	02/11/02	100	-	-	-	100[§]
		Langesund	59°0' N / 9°48' E	09/12/02	100	-	-	-	100[§]
		Inner Skagerrak	58°45' N /10°25' E	04/11/02	100	-	-	-	100[§]
		Jeløya	59°29' N/ 10°37' E	04/11/02	100	-	-	-	100[§]
	Total N				**500**	**100**	-	**-**	**400**
	2003	Høvåg	58°8' N / 8°16' E	03/11/03	100	91	0	0	9
		Kragerø	58°49' N / 9°27' E	17/11/03	100	-	-	-	100[§]
		Langesund	59°0' N / 9°48' E	17/11/03	100	-	-	-	100[§]
		Inner Skagerrak	59°1' N / 10°31' E	19/11/03	100	96	1	0	3
	Total N				**400**	**187**	**1**	**0**	**212**

[§]: Otolith microstructure not examined

Appendix Table A3. Mixture samples: Total number (N) of individuals screened for 9 microsatellite DNA loci and percentage of genotypes scored per locus and sample.

	Skagerrak				**Northern North Sea (Shetland)**	
Locus	Summer 02	Summer 03	Winter 02	Winter 03	2002	2003
N	600	400	500	400	414	310
Cha1017	0.980	0.998	0.996	0.985	0.995	0.835
Cha1020	0.997	0.990	0.992	0.995	0.983	0.981
Cha1027	0.998	1.000	0.978	1.000	0.937	0.961
Cha1202	0.996	1.000	0.994	0.998	0.995	0.990
Cpa101	0.989	0.998	0.988	0.998	0.986	0.994
Cpa111	1.000	0.993	0.994	0.993	0.988	0.994
Cpa112	0.990	0.998	0.978	0.998	0.944	0.926
Cpa113	0.990	1.000	0.986	1.000	0.981	0.961
Cpa114	0.995	1.000	0.992	0.985	0.923	0.890

Appendix Table A4. Estimated composition of mixed feeding aggregation samples. Estimates are provided for the mean, median and the 2.5% and 97.5% percentiles of the frequency distribution of estimates based on the requisite number of MCMC iterations after convergence (Pella and Masuda 2001). Composition of the mixture samples is reported in terms of herring from three broad regions. Region 1 (North Sea autumn spawners, English Channel winter spawners and Norwegian spring spawners). Region 2 (Skagerrak spring spawners). Region 3 [Spring spawners from the Kattegat and western Baltic (Kattegat, Inner Danish Waters and Rügen]. The percentage of juveniles in the mixture samples is indicated

Mixture sample	Year	N	% Juveniles	Region	Mean	Median	2.5%	97.5%
Northern	2002	414	0	1.	0.973	0.980	0.906	0.999
North Sea				2.	0.024	0.015	0.000	0.089
				3.	0.004	0.001	0.000	0.021
	2003	310	0	1.	0.991	0.996	0.948	1.000
				2.	0.006	0.000	0.000	0.048
				3.	0.003	0.001	0.000	0.019
Skagerrak-	2002	600	44.7	1.	0.413	0.413	0.428	0.610
Summer				2.	0.102	0.101	0.000	0.132
				3.	0.486	0.486	0.365	0.526
	2003	400	51.8	1.	0.524	0.525	0.428	0.610
				2.	0.030	0.014	0.000	0.132
				3.	0.447	0.447	0.365	0.526
Skagerrak-	2002	500	10.8	1.	0.476	0.475	0.359	0.600
Autumn/winter				2.	0.402	0.403	0.253	0.543
				3.	0.122	0.119	0.054	0.200

	2003	400	35.2	1.	0.177	0.175	0.066	0.298
				2.	0.698	0.700	0.544	0.835
				3.	0.125	0.124	0.042	0.213

Appendix Table A5. Estimated composition of subsets of mixed feeding aggregation samples classified by Age group and Hatch Season. Age group: Individuals were pooled into two subsets: juveniles, those with no, or one otolith winter ring, and adults, those with two or more otolith winter rings. This analysis was conducted only for the Skagerrak mixture samples as there were no juveniles present in the Northern North Sea mixture samples. Regional estimates are as in Table S4. N.A. indicates that samples did not comprise sufficient numbers of individuals to conduct analysis. **Hatch Season**: individuals in the mixture samples were pooled into three subsets: spring hatched, autumn hatched, and winter hatched. This analysis was conducted for both northern North Sea and Skagerrak mixture samples.

Age group	Sample	Year	N	Region	Mean	Median	2.5%	97.5%
Juveniles (0-1 WR)	**Summer**	2002	310	1	0.708	0.711	0.598	0.800
				2	0.056	0.047	0.000	0.174
				3	0.236	0.235	0.167	0.312
		2003	180	1	0.924	0.929	0.830	0.986
				2	0.015	0.001	0.000	0.100
				3	0.061	0.059	0.009	0.132
	Winter	2002	161	1	0.764	0.772	0.605	0.886
				2	0.078	0.051	0.000	0.295
				3	0.158	0.159	0.042	0.269
		2003	42	1	0.865	0.900	0.571	0.999

				2	0.094	0.031	0.000	0.396
				3	0.041	0.024	0.000	0.166
Adults	**Summer**	2002	289	1	0.234	0.233	0.152	0.316
(2+ WR)				2	0.032	0.011	0.000	0.144
				3	0.735	0.737	0.640	0.815
		2003	220	1	0.179	0.179	0.083	0.277
				2	0.052	0.034	0.000	0.194
				3	0.769	0.773	0.661	0.859
	Winter	2002	297	1	0.272	0.269	0.142	0.415
				2	0.614	0.617	0.453	0.771
				3	0.114	0.112	0.037	0.202
		2003	347	1	0.125	0.122	0.036	0.238
				2	0.718	0.718	0.566	0.862
				3	0.157	0.155	0.064	0.262
Hatch								
season								
Spring	**Shetland**	02/03	22		N.A	NA	N.A	N.A
	Skagerrak-		382	1	0.122	0.122	0.045	0.196
	summer			2	0.022	0.035	0.000	0.133
				3	0.845	0.844	0.767	0.910
	Skagerrak-		287	1	0.169	0.164	0.058	0.307
	winter			2	0.674	0.681	0.487	0.822
				3	0.157	0.155	0.064	0.264
Autumn	**Shetland**		539	1	0.982	0.992	0.917	1.000
				2	0.015	0.003	0.000	0.078
				3	0.004	0.001	0.000	0.019

	Skagerrak-summer		264	1 2 3	0.945 0.022 0.033	0.953 0.003 0.029	0.854 0.000 0.000	0.997 0.119 0.085
	Skagerrak-winter		1		N.A	NA	N.A	N.A
Winter	**Shetland**		118	1 2 3	0.984 0.009 0.008	0.992 0.001 0.003	0.919 0.000 0.000	1.000 0.070 0.042
	Skagerrak-summer		85	1 2 3	0.837 0.107 0.055	0.853 0.079 0.043	0.620 0.000 0.000	0.998 0.348 0.181
	Skagerrak-winter		-		N.A	NA	N.A	N.A

Appendix Table A6. Estimated composition of simulated mixture samples. Simulated mixture samples were constructed with individuals randomly chosen in known proportions from the spawning aggregations sampled in 2002 and tested against three regional baselines based exclusively on allele frequency information for spawners collected in 2003 (See methods). Estimates are provided for the mean, median and the 2.5% and 97.5% percentiles of the frequency distribution of estimates based on the requisite number of MCMC iterations after convergence (Pella and Masuda 2001). Composition of the mixture samples is reported in terms of herring from three broad regions. Region 1 (North Sea autumn spawners, English Channel winter spawners and Norwegian spring spawners). Region 2 (Skagerrak spring spawners). Region 3 [Spring spawners from the Kattegat, Inner Danish Waters and Western Baltic (Rügen)].

Mixture sample	N	Region	Real composition	Mean	Median	2.5%	97.5%
Simulation 1	400	1.	0.33	0.331	0.331	0.225	0.434
		2.	0.33	0.363	0.359	0.243	0.500
		3.	0.33	0.307	0.307	0.223	0.392
Simulation 2	400	1.	0.50	0.485	0.485	0.367	0.603
		2.	0.25	0.238	0.235	0.121	0.365
		3.	0.25	0.278	0.277	0.206	0.350
Simulation 3	400	1.	0.25	0.171	0.169	0.087	0.271
		2.	0.00	0.115	0.115	0.002	0.231
		3.	0.75	0.713	0.714	0.637	0.789
Simulation 4	400	1.	0.75	0.689	0.690	0.579	0.784
		2.	0.05	0.137	0.133	0.049	0.247
		3.	0.20	0.175	0.173	0.121	0.236

Reference list for Electronic Appendix:

Burd AC and Howlett GJ. 1974. Fecundity studies on North Sea herring. Journal du Conseil, 35 (2): 107-120.

Clausen, L.W., Bekkevold, D., Hatfield, E. M. C., and Mosegaard, H. MS. Application and validation of otolith microstructure as stock identifier in mixed Atlantic herring (*Clupea harengus*) stocks in the North Sea and Western Baltic. Cited as unpublished data, paper in preparation for *ICES Journal of Marine Science*.

ICES. 2002. Report of the Planning Group for Herring Surveys Hamburg, Germany 10-14 December 2001. *ICES CM 2002/G:02*.

Nielsen, E. E., Hansen, M. M., Ruzzante, D. E., Meldrup, D., Grønkjær, P. 2003. A hybrid zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual mixture analysis. *Molecular Ecology* **12,** 1497-1508.

Nielsen, E. E., Nielsen, P. H., Meldrup, D., and Hansen, M. M. 2004 Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. *Molecular Ecology* **13**, 585-595.

Van Oosterhout C, Hutchinson WF, Wills PM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes 4:535-538.

Zijlstra JJ. 1969. On the racial structure of North Sea Autumn spawning herring. Journal du Conseil, 33 (1): 67-80.