LACK OF SPATIAL GENETIC SUBSTRUCTURE IN THE 2003-2006 CATCHES OF NA-MINKE WHALES

Hans J. Skaug

Department of Mathematics, University of Bergen, Norway;

ABSTRACT

INTRODUCTION

At its annual meeting in Berlin in 2003 the Scientific Committee conducted an Implementation Review for North Atlantic minke whales (IWC, 2004). Based on genetic (and other) information it was decided to leave the definition of Medium Areas unchanged, while the definitions of Small Areas (SMA) within the eastern Medium Area were changed (Fig. 1). The genetic data available at the time were primarily the Norwegian catches from the years 1997-2002, as contained in Norwegian minke whale DNA-register, but studies based on other markers and samples from earlier years were also reviewed (IWC, 2004). The present paper complements the genetic analyses conducted as part of the Implementation Review, with analyses based on data from the years 2003-2006.

During the Berlin meeting the SMA's were sub-divided into "sites" (Fig. 2). The finer resolution was originally used for the Boundary Rank method, but was eventually adopted for all the analytical methods. The definition of sites remains almost the same in the present analysis, with minor modification made to account for changes the geographical spread of catches. The main conclusion drawn from the 1997-2002 data during the Berlin meeting was that there were differences in mtDNA haplotype frequencies between the Central and Eastern Medium Areas, while the 10 microsatellites showed no such difference. The Eastern Medium Area appeared to be a homogeneous population, with a few poorly understood exceptions. Differences in mtDNA haplotype frequencies were found, especially for females, between animal caught east and west of 28°E along the coast of Finnmark (Hypothesis H14 in Table 6). This finding led to a division of the Barents Sea SMA into a western part (EW) and an eastern part (EB; Fig. 1). An important goal of the present paper is to investigate if the same signal can be found in the 2003-2006 data.

MATERIAL AND METHODS

Origin of samples

Establishment of the Norwegian minke whale DNA register secures that samples (muscle tissues) are taken of each animal caught under the Norwegian catch quota, that the samples are analyzed and that a DNA-profile is established and stored in a database for each whale (Olaisen 1997). The number of samples used in the present paper is 2311 (Table 1), which constitutes 99% of the animals landed in the period 2003-2006. Reasons for exclusion of some animals are: i) sample not obtained from whaling boat, ii) duplication of samples (same profile included twice in the register on the expense of some other profile), iii) problems with interpretation of either mtDNA or microsatellite profile, and iv) inaccuracy in data on geographical position of sample.

Microsatellites
A total of 10 microsatellite loci were used in the analyses (see Table 2 for locus names). Micro variants, i.e. neighboring alleles on the allele-ladder differing by only a single base pair, were merged according to the following scheme:

Locus	**Change**	**# cases**
GATA417	216→217	4
GATA417	221→220	85
GATA417	225→224	12
GATA417	229→228	7
GATA417	231→232	5
GATA028	206→207	25
GATA028	218→219	1

The reason for doing so is that different interpretation of micro variants across years (analysis possibly performed by different personnel) would potentially cause an artificial inhomogeneity in the data.

Test of Hardy-Weinberg equilibrium (against heterozygote deficiency) were performed using the software Genepop (Raymond & Rousset 1995), version 4.0. The same software was used to test for the hypothesis of homogeneity (option 3. 1 in Genepop) and to estimate Fst-values.

mtDNA
Tests of hypothesis about population structure were based on mtDNA haplotype frequencies, using a chi-square test with p-values based on Monte Carlo simulation. Fst-values were calculated using the program Hetero (P. Palsbøll, personal communication).

RESULTS
Table 2 shows p-values for test of Hardy-Weinberg equilibrium for each locus and small area. Applying a Bonferroni correction (or some other multiple testing correction) only locus EV1 in Small Area EW deviates from Hardy-Weinberg equilibrium.

Hardy-Weinberg equilibrium tests were also applied at the 'site' level. Only combinations of level and site with a p-values less than 0.05 are reported (Table 3). Since there are 22 sites, a total of 220 HW-tests have been performed, and hence the expected number of rejections is 11 even when all 220 null hypotheses of HW-equilibrium were true.

Table 4 shows p-values and Fst values for all possible pairwise tests of differences between SMA's, based on microsatellites. The Fst values are all small (less th 0.002) and none of the pairwise tests has a p-value less than 0.05. The results for mtDNA are very much the same (Table 5).

Table 6 shows the hypothesis investigated under the Berlin meeting (Table 4 p. 88 in the Berlin-report). None of the test based on mtDNA were significant. Among the tests based on microsatellites two hypotheses were significant: H13 and H18, both involving site '5W'. On closer inspection there is a single locus (EV37) contributing to this. At EV37 there is an excess of the allele 207 and a corresponding deficiency of the neighboring allele 205. When merging 205 and 207 into a new super-allele the significance goes away. The 205 deficiency applies to all four years.

It was in particular the hypothesis H14 (mtDNA for females) from the Berlin meeting that led to the introduction of a SMA boundary at 28°E. We thus ran the same test (mtDNA for females only) on the

2003-2006 data. The program Hetero (with 10,000 simulations) calculated a negative Fst value (both Hst and Kst) and p-values larger than 0.5 for all three tests (Hst, Kst and chi-square).

The clearest distinction between the Eastern and Central Medium areas was found in mtDNA for females in 2003 (hypothesis H1). We thus ran the same test (mtDNA for females only) on the 2003-2006 data. The program Hetero (with 10,000 simulations) calculated a negative Fst value (both Hst and Kst) and p-values larger than 0.4 for all three tests (Hst, Kst and chi-square).

DISCUSSION AND CONCLUSION

The difference between the Eastern and Central Medium areas (mtDNA) for females in the 1997-2002 data is not found in the current (2003-2006) data. The sample size from the Medium area is only 42 for the current dataset (Table 1; CM), while the sample size was 218 for the 1997-2002 data. Hence the statistical power for the hypothesis test is much lower in the present analysis. However, the observed Fst values between CM and each individual eastern SMA are small (Table 5, last row under females), and hence indicating that there is no difference the Eastern and Central Medium areas.

The difference between east and west of 28°E (hypothesis H14) is not found in the current data. Significant differences in the same region (H13 and H18) is however obtained in the current data, but the estimated Fst values are very low (0.0074 and 3e-04, respectively) representing no real biological signal about two genetically distinct populations being present in the area. Further, the fact that inhomogeneity is caused only by a single locus, and more particularly by a pair of neighbouring alleles at this locus, suggests that some laboratory artefact is causing the effect. It should also be noted that EV27 is out of HW-equilibrium at site 5W (Table 3), but it must be admitted that the situation is not completely understood.

The general conclusion is that 2003-2006 contains very little, if any, signal about spatial genetic structure. One may wonder why the present data contains a lower signal about population structure than the earlier data. One possible explanation is that laboratory techniques have improved, and that the signals in the earlier data were artifacts.

ACKNOWLEDGEMENTS

I am grateful to Gjermund Bøthun and Siri Hartvedt for help with data extraction, and to Per Palsbøll for guidance on the program Hetero.

REFERENCES

IWC (2004). Report of the Sub-Committee on the Revised Management Procedure 2003. J. Cetacean Res. Manage. 4 (Suppl):.

Olaisen, B. (1997). Prposed specifications for a Norwegian DNA database register for minke whales. Int. Whal. Commn SC /49/ NA1: 7 pp.

Raymond M. & Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Heredity, 86:248-249.

Table 1 Norwegian catches of minke whales in the North Atlantic by Small Area for the period 2003-2006.

	EN	EW	EB	ES	CM	TOTAL
2003	94	156	220	147	20	637
2004	81	192	122	109	17	521
2005	6	241	280	95	5	627
2006	30	358	22	116	0	526
TOTAL	211	947	644	467	42	2311

Table 2 P-values for tests of Hardy-Weinberg equilibrium for each microsatellite locus by Small Area. [1]Correpsonding Fis value is -0.013 (Weir and Cockerham variant).

SMA / Locus	EN	EW	EB	ES	CM
GATA098	0.212	0.411	0.469	**0.023**	0.788
GT509	0.743	0.213	0.543	0.872	0.253
EV1	0.915	**0.001**[1]	0.353	0.598	0.056
EV37	0.997	0.205	0.863	0.476	0.698
GT310	0.108	0.176	**0.026**	0.059	0.684
GT211	0.291	0.479	0.306	0.186	0.44
GT575	0.54	0.889	0.27	0.393	0.993
GT023	0.516	0.421	0.333	0.054	0.627
GATA028	0.667	0.944	0.381	0.902	0.654
GATA417	0.497	0.876	0.076	0.812	0.695

Table 3 Summary of sampling sites which deviate from Hardy-Weinberg equilibrium at a given microsatellite locus. Only those cases, out the 220 possible combinations of 'site' and 'locus', with a p-value less than 0.05 are reported. Weir and Cockerham (W&C) variant of Fis is reported.

Locus	Site	# samples	Fis (W&C)	p-value
GATA098	7	189	-0.0182	0.013
GT509	20	17	0.1742	0.017
EV1	9	31	0.1165	0.0179
EV1	18	24	0.1247	0.0183
EV37	5W	128	0.0366	0.024
EV37	18	24	0.2225	0.0097
EV37	20	17	0.1438	0.0394
GT310	4b	252	0.0368	0.0395
GT310	6E	627	0.0711	0.0214
GT310	2	43	0.134	0.0302
GT310	8	87	0.1643	0.0028
GT023	2	43	0.1955	0.0188
GT023	9	31	0.1498	0.0331
GATA417	20	17	0.3361	0.0011

Table 4 Pairwise comparison of SMA's for microsatellites, by sex and combined. Below the diagonal are given Fst-values, while above the diagonal are given p-values for the null hypothesis of homogeneity.

Males and females

	EN	EW	EB	ES	CM
EN		0.5069	0.4468	0.5365	0.933
EW	2e-04		0.6728	0.526	0.8531
EB	2e-04	0		0.3437	0.8416
ES	-3e-04	0	2e-04		0.9081
CM	-0.0015	-0.001	-1e-04	-0.0011	

Males

	EN	EW	EB	ES	CM
EN		0.4977	0.2666	0.5933	0.9291
EW	3e-04		0.4608	0.1891	0.9379
EB	7e-04	-2e-04		0.2371	0.8653
ES	-0.0018	0	-7e-04		0.9034
CM	-0.0021	-0.0038	-0.0028	-0.0026	

Females

	EN	EW	EB	ES	CM
EN		0.5423	0.2525	0.2662	0.9758
EW	-1e-04		0.7627	0.7742	0.9714
EB	-1e-04	-3e-04		0.3649	0.9556
ES	-2e-04	-3e-04	2e-04		0.9886
CM	-0.0021	-9e-04	3e-04	-0.0015	

Table 5 Pairwise comparison of Small Areas for mtDNA, by sex and combined. Below the diagonal are given Fst-values (Hst), while above the diagonal are given p-values for the null hypothesis of homogeneity.

Males and females

	EN	EW	EB	ES	CM
EN		0.1527	0.423	0.3249	0.7503
EW	3e-04		0.3835	0.7822	0.7498
EB	0	0		0.745	0.7358
ES	2e-04	-2e-04	-3e-04		0.8455
CM	-0.0013	-3e-04	-4e-04	-7e-04	

Males

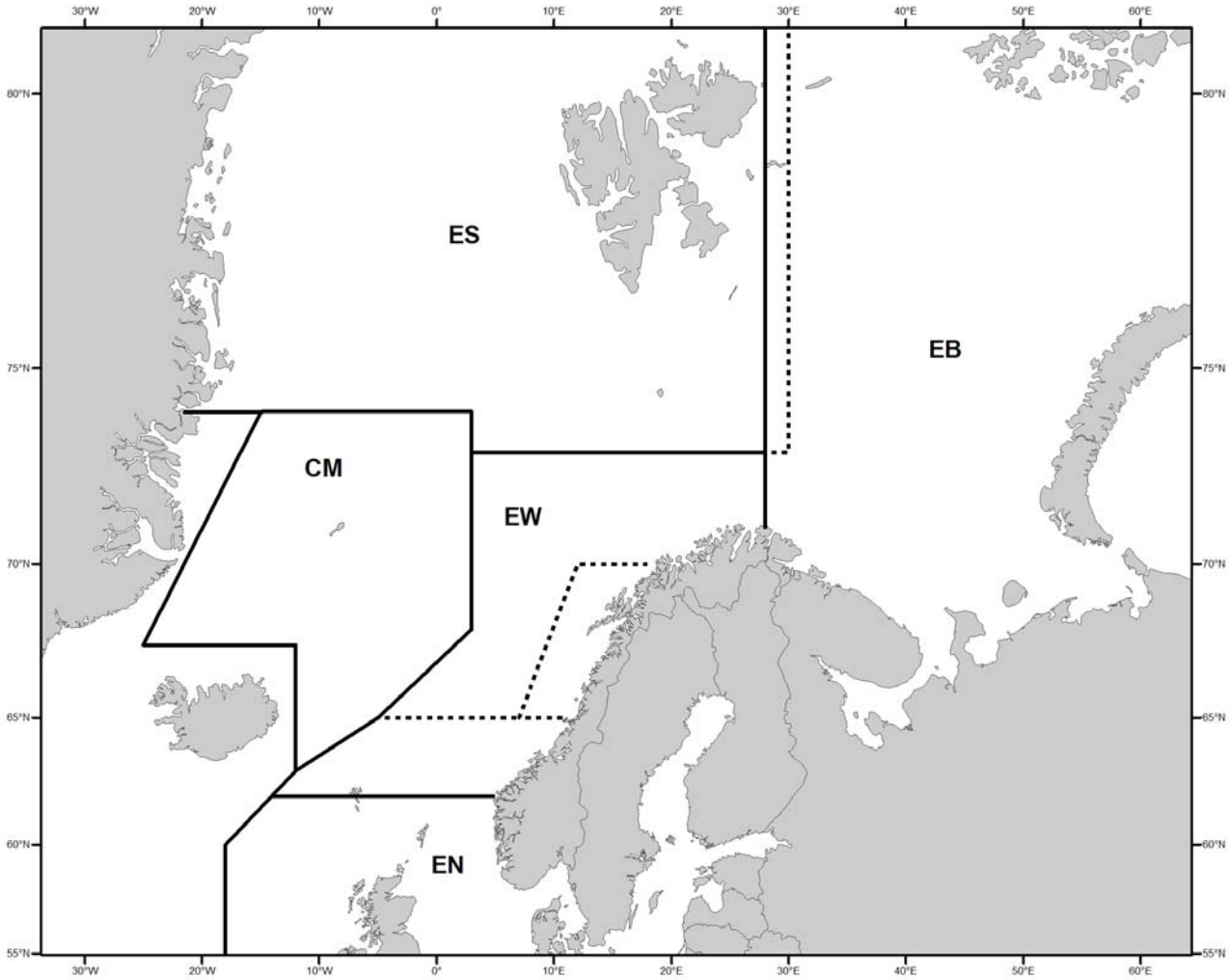
	EN	EW	EB	ES	CM
EN		0.5915	0.9113	0.963	0.4403
EW	-3e-04		0.3939	0.808	0.2165
EB	-0.0019	0		0.2763	0.2251
ES	-0.0049	-6e-04	9e-04		0.776
CM	-5e-04	6e-04	0.0017	-0.0148	

Females

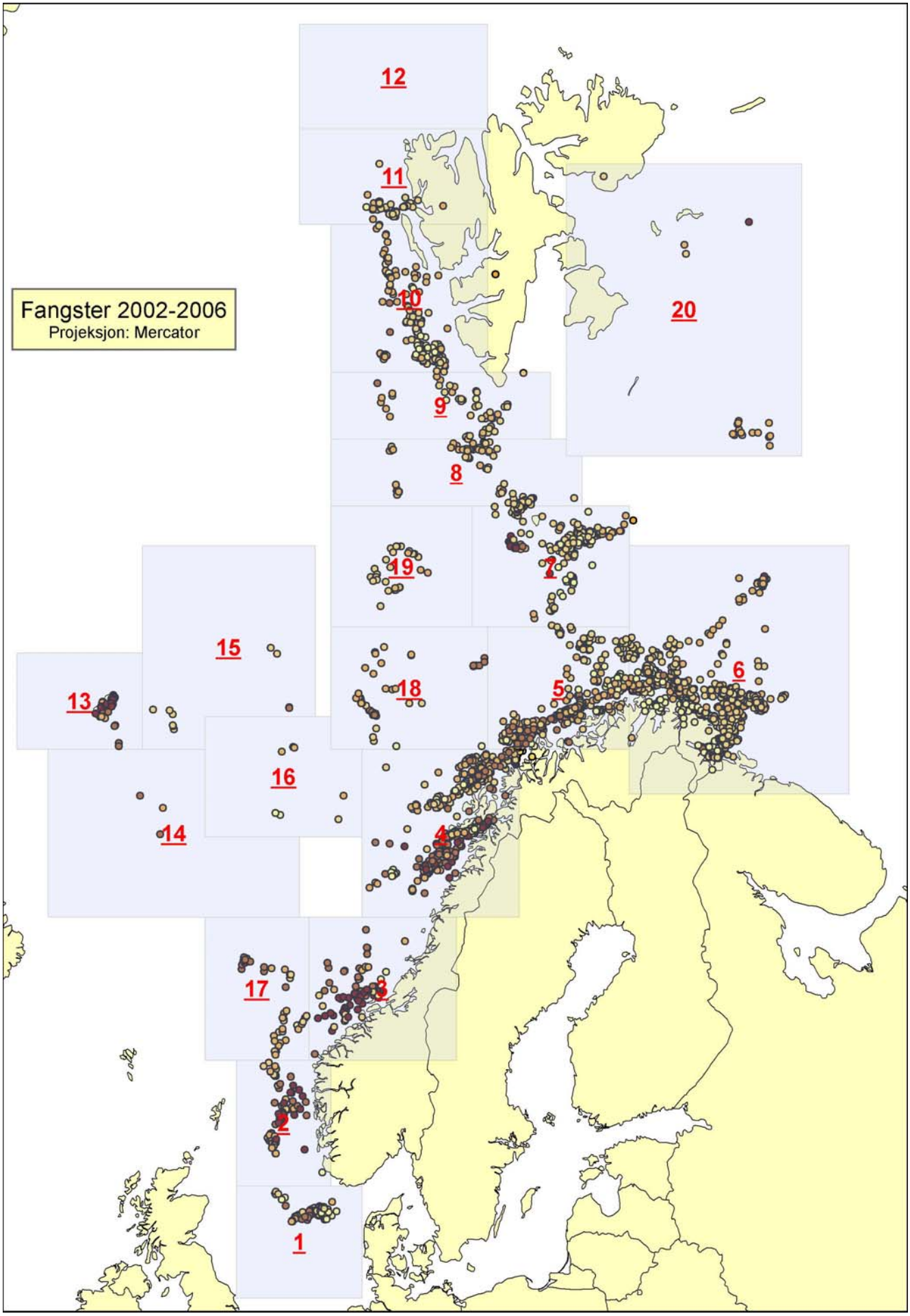
	EN	EW	EB	ES	CM
EN		0.5846	0.4497	0.5456	0.5326
EW	-3e-04		0.8498	0.361	0.3973
EB	-1e-04	-4e-04		0.84	0.4036
ES	-2e-04	0	-4e-04		0.3605
CM	-8e-04	0	0	1e-04	

Table 6 Test of hypotheses about combined small areas. In all cases the null hypothesis is that of no population substructure (populations 1,2,3 constitute a single population). Sample size *n* is given in parenthesis.

				Microsatellites		**mtDNA**
	Area 1 (*n*)	**Area 2 (*n*)**	**Area 3 (*n*)**	**F_{st}**	**p-value**	**p-value**
H1	CM (42)	EN,EW,EB,ES (2269)		4e-04	0.8463	0.638
H2	CM (42)	EW,EB,ES (2058)		0.0044	0.8665	0.6974
H3	EW (947)	EN (211)		0.0037	0.6439	0.4444
H4	EN (211)	EW,EB,ES (2058)		0.0037	0.5104	0.6337
H5	EN,EW (1158)	EB,ES (1111)		0.0037	0.6418	0.2331
H6	EN (211)	EW,EB,ES,CM (2100)		0.0043	0.5201	0.6572
H7	1,2 (211)	6W,6E (776)	7-12 (465)	0.0055	0.5553	0.2763
H8	6W,6E (776)	7-12 (465)		0.0068	0.6372	0.7403
H9	14 (5)	13,15,16 (37)		0.0422	0.7007	0.8256
H10	1,2 (211)	3 (60)		-0.0012	0.3961	0.3157
H11	1,2 (211)	3,4 (316)		0.0037	0.0894	0.1112
H12	1,2 (211)	3,4,4b,5W (696)		4e-04	0.3989	0.4824
H13	5W (128)	5E,6W,6E (836)		0.0074	**0.0028**	0.1124
H14	5,6W (149)	6E (627)		4e-04	0.2483	0.8322
H15	5W (128)	5E (60)		-0.002	0.4544	0.6511
H16	5E (60)	6W (149)		-0.0032	0.4838	0.1465
H17	6W (149)	6E (627)		0.0097	0.2483	0.8318
H18	4 (256)	5W (128)		3e-04	**0.0027**	0.6099



Figur 1 Small Area boundaries adopted at the 2003 meeting in Berlin (solid line). The pre-2003 boundaries are shown as dashed lines where differerent from the current ones.



Figur 2 Positions of catches and definitions of 'sites' (1-20).