

Origin and route of establishment of the invasive Pacific oyster *Crassostrea gigas* in Scandinavia

Ellika Faust*, Carl André, Sara Meurling, Judith Kochmann, Henrik Christiansen,
Lasse Fast Jensen, Grégory Charrier, Ane T. Laugen, Åsa Strand

*Corresponding author: ellika.faust@gmail.com

Marine Ecology Progress Series 575: 95–105 (2017)

Table S1 Overview of the genetic diversity. Observed and unbiased expected heterozygosity (H_o , uH_e) were calculated in GenAlEx 6.501 (Peakall & Smouse 2006, 2012), and number of alleles (N_a) and allelic richness (A_R) were calculated in FSTAT (Goudet 2002).

		France		Netherlands		Ireland			Denmark		Sweden						Norway			
AMY	Na	32	34	27	31	17	21	20	20	21	24	25	19	20	20	28	23	24	20	14
	Ar	21	21	16	20	11	13	14	14	15	16	14	14	14	14	17	15	16	14	14
	Ho	0.98	0.86	0.83	0.88	0.98	0.88	0.86	0.90	0.82	0.92	0.88	0.85	0.92	0.96	0.92	0.92	0.95	0.91	1.00
	uHe	0.96	0.96	0.93	0.96	0.87	0.89	0.91	0.92	0.92	0.92	0.92	0.92	0.92	0.90	0.92	0.92	0.90	0.93	0.90
L48	Na	41	45	26	40	19	18	28	24	27	25	23	25	28	28	32	29	28	25	10
	Ar	24	23	16	23	13	13	18	16	17	15	15	16	17	18	19	18	17	17	10
	Ho	0.93	0.90	0.89	0.96	0.92	0.90	0.87	0.94	0.96	0.90	0.94	0.96	0.92	0.96	0.94	0.92	0.93	0.98	0.95
	uHe	0.98	0.97	0.93	0.97	0.90	0.92	0.94	0.94	0.94	0.92	0.93	0.93	0.93	0.95	0.95	0.95	0.94	0.94	0.87
CGE09	Na	21	19	14	18	12	10	12	11	11	13	12	10	10	13	14	12	12	10	8
	Ar	14	13	11	13	9	8	10	9	9	10	9	9	9	11	10	9	10	8	8
	Ho	0.83	0.90	0.88	0.88	0.81	0.78	0.82	0.76	0.84	0.92	0.85	0.92	0.80	0.88	0.86	0.74	0.93	0.87	1.00
	uHe	0.91	0.89	0.89	0.92	0.83	0.84	0.88	0.85	0.86	0.86	0.84	0.86	0.83	0.89	0.87	0.84	0.89	0.80	0.89
Cg108	Na	36	34	26	34	16	16	19	17	23	23	20	21	21	25	24	25	29	17	15
	Ar	22	20	16	22	12	13	13	13	15	15	15	14	15	16	17	17	17	13	14
	Ho	0.96	0.78	0.85	0.90	0.85	0.88	0.84	0.82	0.94	0.88	0.90	0.86	0.84	0.94	0.84	0.92	0.85	0.85	0.89
	uHe	0.97	0.95	0.92	0.97	0.90	0.91	0.90	0.92	0.93	0.94	0.94	0.92	0.93	0.94	0.95	0.95	0.94	0.91	0.91
L10	Na	33	39	25	37	19	20	20	22	22	25	24	26	27	22	24	29	28	23	12
	Ar	22	21	16	22	14	14	14	15	16	17	17	17	18	15	17	19	19	16	11
	Ho	0.96	0.90	0.85	0.96	0.80	0.92	0.94	0.94	0.92	0.90	0.89	0.94	0.90	0.96	0.96	0.94	0.98	0.98	0.89
	uHe	0.97	0.96	0.92	0.97	0.92	0.91	0.92	0.93	0.93	0.94	0.95	0.94	0.95	0.94	0.94	0.95	0.96	0.94	0.88
Cgsili44	Na	17	18	17	17	11	14	15	14	14	16	15	14	16	13	16	16	16	16	9
	Ar	14	14	12	14	11	11	12	11	11	13	12	12	13	12	13	13	12	12	9
	Ho	0.86	0.80	0.87	0.88	0.94	0.86	0.88	0.84	0.80	0.91	0.85	0.74	0.89	0.89	0.91	0.93	0.93	0.85	0.71
	uHe	0.93	0.91	0.87	0.93	0.90	0.89	0.92	0.90	0.90	0.91	0.90	0.88	0.91	0.91	0.92	0.93	0.91	0.88	0.85

H_o = Observed Heterozygosity = No. of Hets / N. uH_e = Unbiased Expected Heterozygosity = $(2N / (2N-1)) * 1 - \sum p_i^2$. N_a = No. of alleles. A_R = Allelic richness.
Swedish size classes are displayed as: 1 = larger individuals (91-239 mm) and 2 = smaller individuals (29-81 mm).

*Aquaculture samples.

Table S2 Each locus in each sample was tested for deviations from Hardy-Weinberg (HW) proportions by calculating F_{IS} and p-values for the exact HW test in Genepop 4.3 (Rousset 2008). P-values significant after FDR correction are marked in **bold**. Equivalent F_{IS} values are also marked in **bold**.

	P-values for Hardy Weinberg Probability test							FIS						
	AMY	L48	CGE09	Cg108	L10	Cgsili44	Over All Loci	AMY	L48	CGE09	Cg108	L10	Cgsili44	Over All Loci
France Marennes-Oléron*	0.830	0.000	0.094	0.884	0.170	0.353	0.0000	-0.015	0.044	0.091	0.012	0.015	0.070	0.0353
France Marennes-Oléron	0.018	0.178	0.988	0.001	0.010	0.004	0.0000	0.112	0.075	-0.006	0.185	0.065	0.124	0.093
Netherlands Oosterschelde*	0.009	0.068	0.382	0.000	0.002	0.297	0.0000	0.106	0.038	0.012	0.078	0.080	0.010	0.0543
Netherlands Oosterschelde	0.000	0.322	0.261	0.020	0.446	0.265	0.0011	0.092	0.016	0.042	0.076	0.012	0.053	0.0487
Ireland Lough Foyle*	0.015	0.028	0.010	0.000	0.000	0.000	0.0000	-0.123	-0.014	0.022	0.059	0.135	-0.049	0.0046
Ireland Lough Foyle	0.589	0.103	0.116	0.071	0.696	0.491	0.1368	0.010	0.024	0.068	0.033	-0.009	0.030	0.0252
Ireland Lough Swilly	0.454	0.060	0.539	0.236	0.980	0.045	0.1289	0.054	0.074	0.070	0.068	-0.016	0.047	0.0493
Denmark Wadden Sea	0.584	0.727	0.246	0.063	0.772	0.093	0.2248	0.018	-0.005	0.109	0.113	-0.011	0.064	0.0466
Denmark Limfjorden	0.064	0.875	0.256	0.289	0.299	0.260	0.1872	0.115	-0.020	0.024	-0.007	0.012	0.103	0.0365
Sweden Smalsundet (1)	0.676	0.157	0.615	0.121	0.343	0.110	0.1808	-0.002	0.028	-0.074	0.060	0.043	0.002	0.0107
Sweden Smalsundet (2)	0.321	0.112	0.956	0.659	0.026	0.128	0.0882	-0.037	-0.014	0.013	0.000	-0.022	0.032	0.0276
Sweden Furulund (1)	0.137	0.430	0.605	0.658	0.430	0.020	0.1487	0.041	-0.014	-0.015	0.043	0.056	0.056	0.0318
Sweden Krokesundet (1)	0.892	0.780	0.794	0.025	0.075	0.124	0.1170	0.074	-0.029	-0.065	0.064	0.003	0.160	0.0366
Sweden Krokesundet (2)	0.356	0.518	0.747	0.176	0.761	0.255	0.5528	0.001	0.017	0.012	0.117	-0.018	0.009	-0.0056
Sweden Svallhagen (1)	0.580	0.157	0.625	0.032	0.541	0.780	0.2795	-0.018	0.018	0.038	0.100	0.052	0.026	0.0239
Sweden Svallhagen (2)	0.485	0.150	0.035	0.003	0.770	0.405	0.0104	-0.021	0.028	0.124	0.028	0.014	0.001	0.0284
Norway Tromlingene	0.129	0.446	0.685	0.117	0.957	0.359	0.3773	-0.028	0.011	-0.055	0.098	-0.024	-0.020	-0.0018
Norway Hui	0.428	0.659	0.709	0.619	0.724	0.064	0.5875	-0.016	-0.045	-0.094	0.070	-0.040	0.036	-0.0144
Norway Espevik	0.648	0.980	0.356	0.548	0.251	0.160	0.5624	-0.084	-0.087	-0.125	0.019	-0.018	0.176	-0.0245

Table S3 Matrix of pairwise F_{ST} (θ) according to Weir and Cockerham (1984) calculated in GENETIX 4.05 (Belkhir et al. 1996-2004). The significance ($p < 0.05$) of the F_{ST} values were tested using 9999 permutations, and were corrected for multiple testing using the False Discovery Rate (FDR) correction (Benjamini & Hochberg 1995).

		France		Netherlands		Ireland			Denmark		Sweden						Norway				
		Marennes-Oléron#	Marennes-Oléron	Oosterschelde#	Oosterschelde	Lough Foyle#	Lough Foyle	Lough Swilly	Wadden Sea	Limfjorden	Smalsundet 1	Furulund 1	Krokesundet 1	Svallhagen 1	Smalsundet 2	Krokesundet 2	Svallhagen 2	Tromlingene	Hui	Espevik	
France	Marennes-Oléron [‡]		0.0004	0.0171***	-0.0016	0.0334***	0.0277***	0.0204***	0.0133***	0.0153***	0.0136***	0.0138***	0.0154***	0.0144***	0.0115***	0.0089***	0.0123***	0.0079***	0.0269***	0.0330***	
	Marennes-Oléron			0.0217***	0.0027	0.0358***	0.0364***	0.0260***	0.0186***	0.0210***	0.0191***	0.0216***	0.0216***	0.0197***	0.0153***	0.0151***	0.0193***	0.0147***	0.0358***	0.0415***	
Netherlands	Oosterschelde [‡]				0.0150***	0.0414***	0.0372***	0.0321***	0.0269***	0.0266***	0.0259***	0.0253***	0.0270***	0.0232***	0.0245***	0.0211***	0.0216***	0.0178***	0.0296***	0.0460***	
	Oosterschelde					0.0315***	0.0290***	0.0199***	0.0167***	0.0160***	0.0158***	0.0144***	0.0181***	0.0168***	0.0112***	0.0097***	0.0143***	0.0085***	0.0266***	0.0289***	
Ireland	Lough Foyle [‡]						0.0260***	0.0223***	0.0375***	0.0335***	0.0324***	0.0332***	0.0338***	0.0358***	0.0313***	0.0346***	0.0312***	0.0298***	0.0492***	0.0405***	
	Lough Foyle							0.0111***	0.0159***	0.0112***	0.0165***	0.0139***	0.0093***	0.0155***	0.0146***	0.0185***	0.0173***	0.0118***	0.0235***	0.0198***	
	Lough Swilly								0.0117***	0.0090***	0.0149***	0.0141***	0.0118***	0.0176***	0.0107***	0.0162***	0.0137***	0.0065***	0.0268***	0.0223**	
Denmark	Wadden Sea									-0.0034	0.0001	-0.0020	0.0003	0.0007	0.0012	0.0006	-0.0012	-0.0015	0.0108***	0.0239***	
	Limfjorden										-0.0011	-0.0011	-0.0011	0.0002	0.0001	0.0010	-0.0012	-0.0023	0.0071**	0.0170***	
Sweden	Smalsundet(1)											-0.0023	0.0031	-0.0003	0.0023	-0.0004	-0.0026	-0.0014	0.0053**	0.0209***	
	Furulund(1)												0.0021	-0.0021	0.0006	0.0022	-0.0017	0.0010	0.0124***	0.0186***	
	Krokesundet(1)													0.0019	0.0009	0.0040*	0.0035	-0.0008	0.0070**	0.0157***	
	Svallhagen(1)														0.0016	0.0004	-0.0007	0.0012	0.0092***	0.0222***	
	Smalsundet(2)															0.0015	0.0016	-0.0003	0.0067**	0.0226***	
	Krokesundet(2)																-0.0021	0.0002	0.0065**	0.0248***	
	Svallhagen(2)																	-0.0015	0.0051*	0.0257***	
Norway	Tromlingene																		0.0073**	0.0218***	
	Hui																				0.0317***
	Espevik																				

Swedish size classes are displayed as; 1 = larger individuals (91-239 mm) and 2 = smaller individuals (29-81 mm). Significance after FDR correction displayed as * $q < 0.05$ ** $q < 0,01$ and *** $q < 0,001$. Swedish size cohorts are displayed in brackets: (1) larger individuals and (2) smaller individuals [‡] Aquaculture samples.

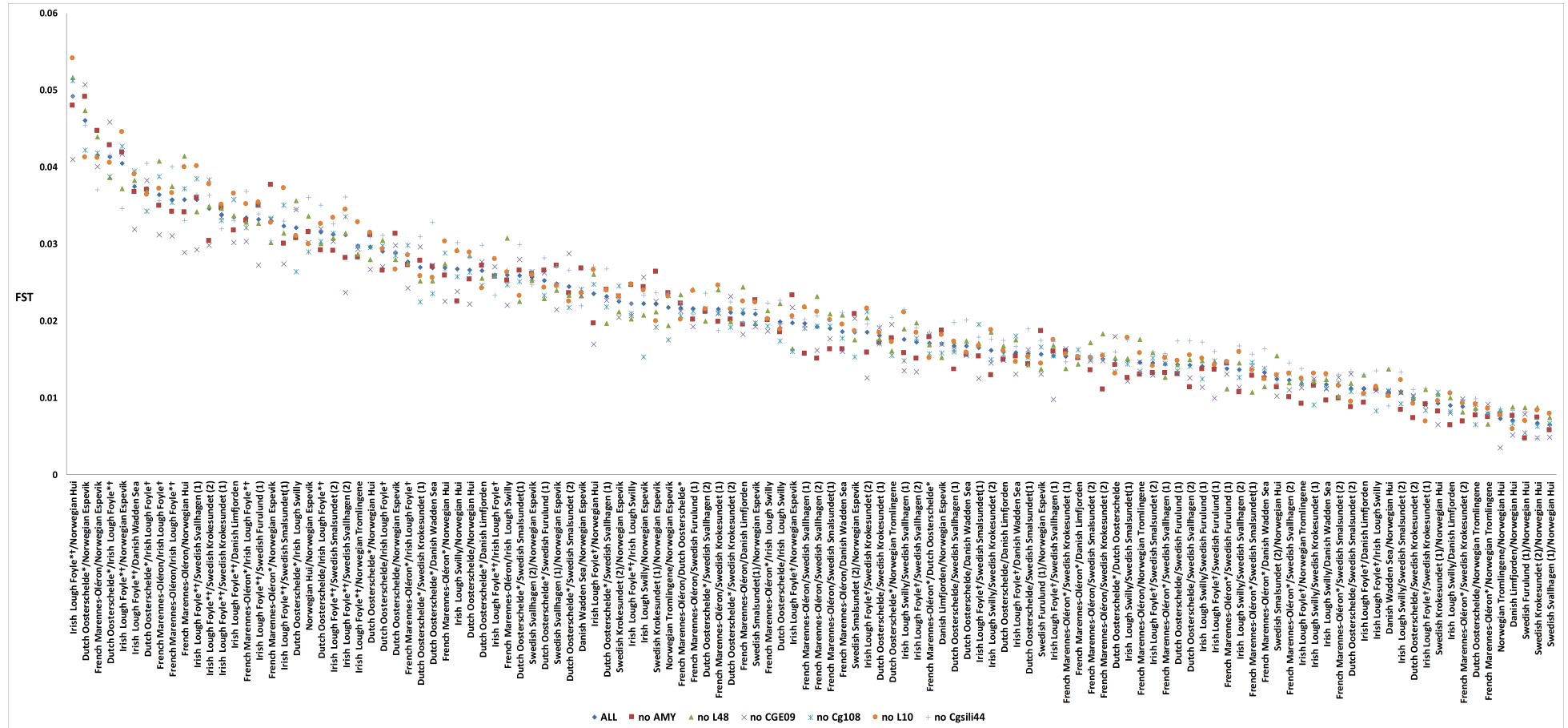


Figure S1 Jack-knife resampling over loci for Pair-wise F_{ST} estimates, where each locus was systematically omitted. Swedish size classes are displayed in brackets: (1) larger individuals and (2) smaller individuals. *Aquaculture samples.

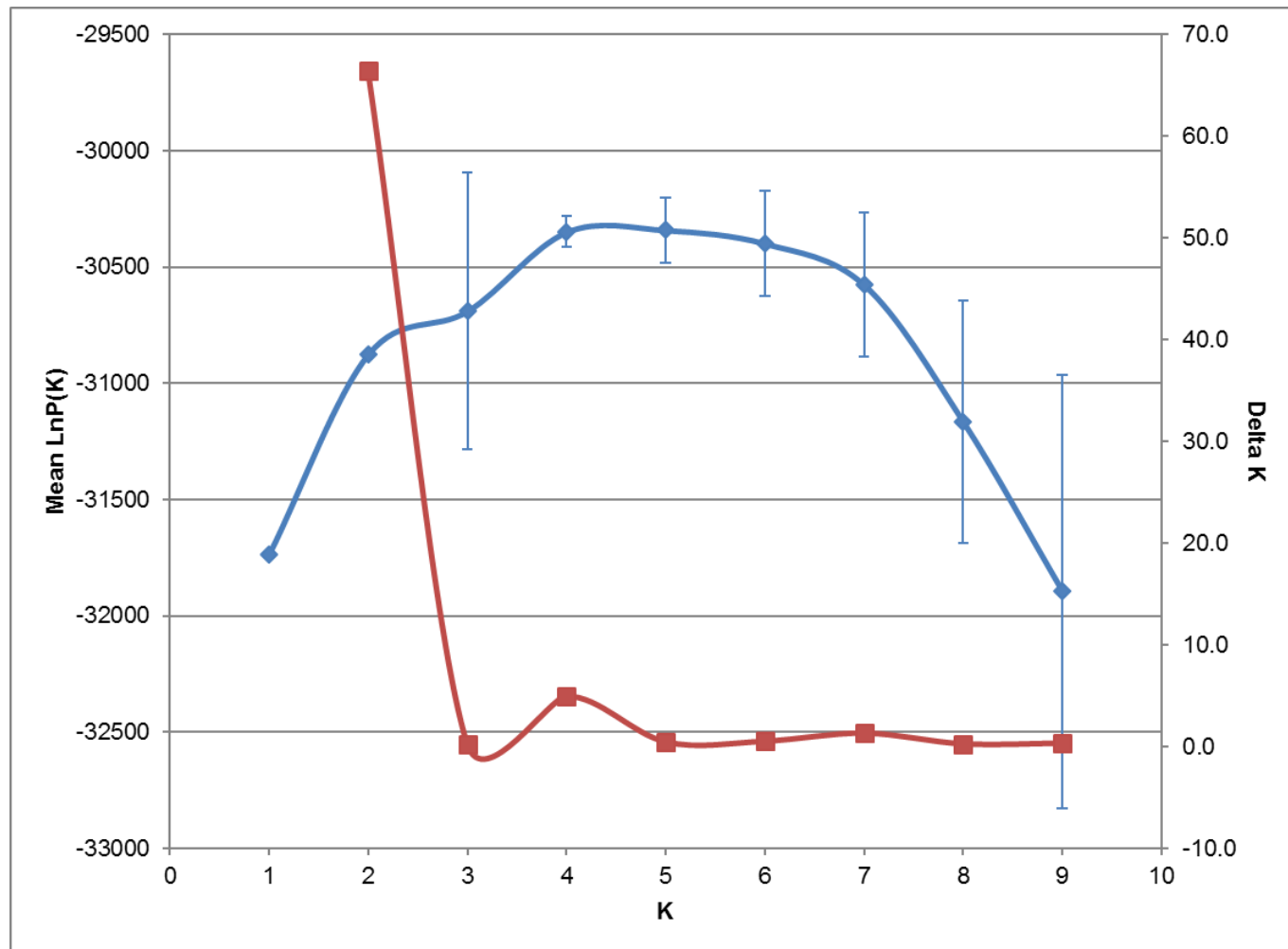


Figure S2 A visualisation of mean $\text{LnP}(K) \pm \text{SD}$ and delta K for K 1-9 clusters run 20 times in STRUCTURE (Pritchard et al. 2000). Values calculated using Structure Harvester, <http://taylor0.biology.ucla.edu/structureHarvester> (Earl & Von Holdt 2012).

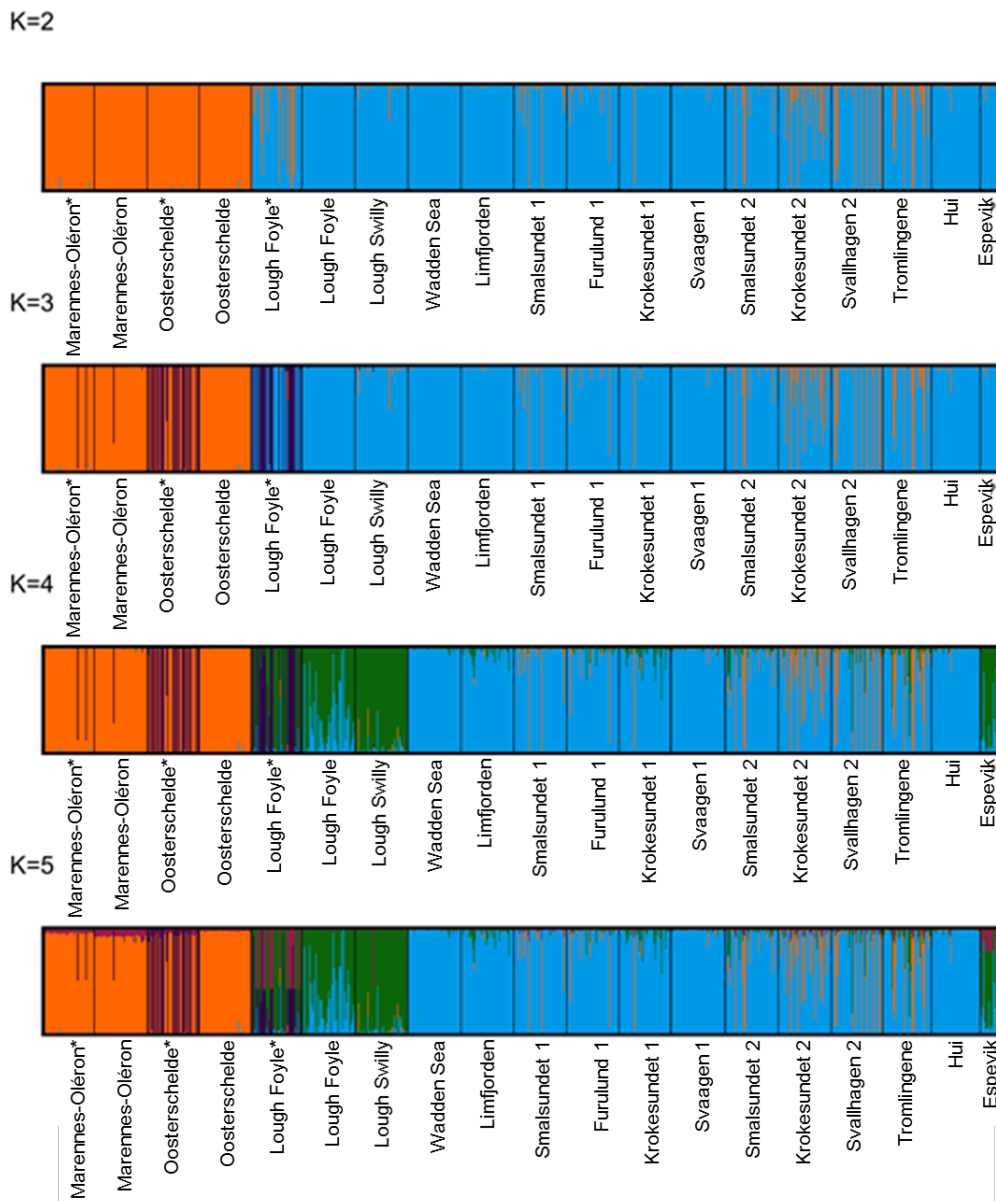


Figure S3 Genetic population structure of *Crassostrea gigas* based on six microsatellite markers. STRUCTURE output for all samples visualised with CLUMPAK (Kopelman et al. 2015) showing major modes for selected $K=2-5$. Every vertical line represents one individual and the colour shows the proportion of each individual assigned to different each of the four genetic clusters. Swedish size classes are displayed as; 1 = larger individuals (91-239 mm) and 2 = smaller individuals (29-81 mm). *Aquaculture samples

References

- Belkhir, K. et al., 2004. GENETIX 4.05, logiciel sous Windows™ pour la génétique des populations. Laboratoire Genome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France.
- Benjamini, Y. & Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), pp.289 – 300.
- Earl, D.A. & Von Holdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), pp.359–361.
- Goudet, J., 2002. FSTAT, a program to estimate and test gene diversities and differentiation statistics from codominant genetic markers (version 2.9.3.2.). Available at: <http://www2.unil.ch/popgen/softwares/fstat.htm> [Accessed May 26, 2015].
- Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour* 15:1179–1191.
- Peakall, R. & Smouse, P.E., 2012. GenALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28(19), pp.2537–2539.
- Peakall, R. & Smouse, P.E., 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), pp.288–295.
- Pritchard, J.K., Stephens, M. & Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2), pp.945–959.
- Rousset, F., 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8(1), pp.103–106.
- Weir, B.S. & Cockerham, C.C., 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, 38(6), pp.1358–1370.