

MOLECULAR ECOLOGY RESOURCES

Supplemental Information for:

Management and conservation implications for cryptic population sub-structure for two commercially exploited fishes (*Merluccius* spp.) in southern Africa

Sarah Forde¹, Sophie von der Heyden², Alan Le Moan³, Erica S Nielsen⁴, Deon Durholtz⁵, Paulus Kainge⁶, Johannes N Kathena⁶, Marek R Lipinski⁷, Hilka ON Ndjaula⁸, Conrad A Matthee², and Romina Henriques¹

Affiliations: ¹Marine Genomics Group, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa; ²Evolutionary Genomics Group, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Stellenbosch 7602 South Africa; ³Sven Loven Centre for Marine Sciences, University of Gothenburg, Box 461 Goteborg 40530, Sweden; ⁴University of California – Davis, United States of America; ⁵Department of Forestry, Fisheries and Environment, Cape Town, South Africa; ⁶National Marine Information and Research Centre, Ministry of Fisheries and Marine Resources, Swakopmund, Namibia; ⁷Department of Ichthyology and Fisheries Science, Rhodes University, Makhanda 6140, South Africa; ⁸Sam Nujoma Marine and Coastal Resources Research Centre, University of Namibia, P.O. Box 642, Henties Bay, Namibia.

Table S1: Summary for the ezRAD Pool-Seq experiment for the species *Merluccius capensis* (C) and *M. paradoxus* (P), including the pool sequenced (NN – north Namibia, CN – Central Namibia, WC – north west coast South Africa, WC2 – south west coast South Africa, SW – southwest coast South Africa); the average latitude and longitude coordinates of the sampling points (Average latitude/longitude of sampling), average depth of sampling points (Average depth of sampling (m)); the number of raw reads (Raw read pair); the retained number of read pairs after trimming (Retained read pairs); the average sequencing depth (Average sequencing coverage); the number of reads that mapped to mitochondrial DNA (Reads mapped to mtDNA); the number of reads that mapped to the reference genome with a quality of 20 or above (Reads mapped to reference genome (q>20)); and the subsampling proportion used to standardize the number of reads per pool (Sub-sampling fraction).

Species	Pool	Number of individuals	Average latitude / longitude of sampling	Average depth of sampling (m)	Raw read pairs	Retained read pairs	Average sequencing coverage	Reads mapped to mtDNA	Reads mapped to reference genome (q>20)	Sub-sampling fraction
<i>M. capensis</i>	CNN	37	-18.00 / 11.42	310.5	3 504 836	3 344 273	17.96	11 447	1 999 406	0.312
	CCN	35	-26.00 / 14.92	300.5	3 280 405	3 151 717	16.12	10 620	624 081	1.000
	CWC	40	-30.18 / 15.45	233	2 863 204	2 594 346	14.93	6 358	1 456 415	0.429
	CWC2	20	-32.18 / 16.99	262.5	3 309 290	2 750 746	18.24	6 373	1 887 818	0.331
<i>M. paradoxus</i>	PNN	40	-19.00 / 13.52	507	3 373 994	3 287 878	30.77	16 964	1 022 282	1.000
	PCN	35	-24.67 / 13.52	443	3 420 308	2 444 810	23.62	9 601	1 871 684	0.546
	PWC	36	-29.97 / 15.08	393	2 948 743	2 777 036	25.43	12 163	1 788 824	0.571
	PWC2	37	-32.81 / 16.99	321	3 843 477	3 221 952	31.08	13 753	1 541 330	0.663
	PSW	38	-35.27 / 18.71	518	3 445 704	3 022 846	21.98	6 994	2 050 467	0.499

MOLECULAR ECOLOGY RESOURCES

Table S2: Average genetic diversity metrics per species and pool, including nucleotide diversity (π), Watterson's theta (θ) and Tajima's D (D), calculated in 500 bp windows for each species and pool, as well as the number of private SNPs per pool. Pool names as per Table S1.

Species	Pool	π	θ	D	Number of private SNPs
<i>M. capensis</i>	CNN	0.0154	0.0192	-0.8394	914
	CCN	0.0120	0.0166	-0.9514	522
	CWC	0.0133	0.0173	-0.8177	616
	CWC2	0.0203	0.0252	-0.8268	2923
	Total	0.0154	0.0192	-0.8394	-
<i>M. paradoxus</i>	PNN	0.0214	0.0259	-0.9134	2237
	PCN	0.0195	0.0237	-0.8159	1181
	PWC	0.0208	0.0253	-0.7641	1541
	PWC2	0.0204	0.0255	-0.9828	2371
	PSW	0.0204	0.0260	-0.8063	3493
	Total	0.0214	0.0259	-0.9134	-

Table S3: Pairwise F_{ST} comparison values, obtained with poolstat, for *M. capensis* pools for the final SNP dataset, after accounting for batch effect SNPs, using the exact pool size (2 300 SNPs). *significantly different from 0. Pool names as per Table S1.

Pool	CNN	CCN	CWC	CWC2
CNN	x			
CCN	-0.005	x		
CWC	-0.002	-0.007	x	
CWC2	0.013*	-0.002	0.002	x

Table S4: Pairwise F_{ST} comparison values, obtained with poolstat, for *M. paradoxus* pools for the final SNP dataset, after accounting for batch effect SNPs, using the exact pool size (below diagonal – 2 250 SNPs). *significantly different from 0. Pool names as per Table S1.

Pool	PNN	PCN	PWC	PWC2	PSW
PNN	x				
PCN	-0.006	x			
PWC	-0.008	-0.013	x		
PWC2	-0.011	-0.005	-0.007	x	
PSW	-0.004	-0.004	-0.010	0.001	x

Table S5: Results of the demographic inferences inferred from *moments*. The table shows the species analysed (Species), the pairs of populations used (Pop1 and Pop2), the demographic models (panmixia, PAM, or isolation-with-migration, IM), the likelihood and the Akaike Information Criterion (AIC) of the model, the estimates of ancestral population-mutation rate (θ), and the parameter of the model (given relative to θ). For the IM model, the parameters are the effective population sizes of Pop 1 (N1) and Pop2 (N2), the time of split of the two populations (T) and the number of effective migrants from population 1 to population 2 (M1>2) or from population 2 to 1 (M2>1). The parameters of the PAM model are the effective population size of the current panmictic population (N1) and the time at which the ancestral population size has changed (T). Pool names as per Table S1.

Species	Pool 1	Pool 2	Model	likelihood	AIC	θ	N1	N2	T	M1>2	M2>1
	CNN	CCN	PAM	-279.368	562.736	433.636	5.422		0.327		

MOLECULAR ECOLOGY RESOURCES

<i>M. capensis</i>	CNN	CWC	IM	-276.605	563.211	432.059	5.909	4.586	0.329	30.900	2.593	
			IM	-281.641	573.283	374.061	5.517	0.648	0.381	6.596	56.284	
	CNN	CWC2	PAM	-293.816	591.633	379.450	5.910		0.411			
			IM	-350.225	710.450	430.051	7.111	0.098	0.328	8.321	115.162	
	CCN	CWC	PAM	-436.066	876.132	447.359	6.163		0.321			
			IM	-261.355	532.710	357.996	5.574	1.021	0.394	9.234	49.404	
	CCN	CWC2	PAM	-268.207	540.415	362.002	5.939		0.440			
			IM	-312.142	634.283	428.993	5.974	0.343	0.324	5.640	50.084	
	CWC	CWC2	PAM	-358.845	721.690	438.707	6.0787		0.334			
			IM	-314.902	639.805	380.529	6.005	0.265	0.399	6.397	70.794	
			PAM	-350.595	705.189	390.744	6.039		0.412			
			IM	-218.615	447.229	174.976	15.529	10.764	0.794	10.114	0.012	
	<i>M. paradoxus</i>	PNN	PCN	PAM	-223.012	450.023	175.690	11.438		0.749		
				IM	-209.081	428.161	173.686	12.093	11.186	0.815	10.283	0.008
PNN		PWC	PAM	-213.522	431.045	174.418	11.884		0.774			
			PAM	-206.503	417.006	177.102	10.396		0.774			
PNN		PWC2	IM	-205.865	421.729	175.494	4.217	8.984	0.698	34.424	5.743	
			IM	-226.754	463.508	179.383	2.624	10.951	0.741	19.181	0.928	
PNN		PSW	PAM	-236.056	476.113	179.704	11.724		0.765			
			PAM	-204.689	413.379	173.621	12.847		0.792			
PCN		PWC	IM	-204.699	419.397	172.511	11.091	68.866	0.764	2.645	23.721	
			IM	-222.257	454.515	176.100	1.781	11.705	0.688	32.930	3.490	
PCN		PWC2	PAM	-227.785	459.570	178.414	11.875		0.747			
			IM	-222.381	454.761	174.483	1.230	18.239	0.694	31.283	7.626	
PCN		PSW	PAM	-231.307	466.614	177.012	11.334		0.785			
			IM	-214.581	439.162	173.315	10.381	16.910	0.837	0.433	9.814	
PWC		PWC2	PAM	-218.713	441.427	174.934	11.831		0.792			
			IM	-213.585	437.169	174.067	10.468	99.796	0.883	0.298	9.708	
PWC		PSW	PAM	-216.763	437.526	176.123	11.409		0.814			
			IM	-241.299	492.599	180.295	4.275	9.296	0.746	7.917	3.224	
PWC2		PSW	PAM	-256.110	516.220	181.661	12.775		0.785			
			IM	-218.615	447.229	174.976	15.529	10.764	0.794	10.114	0.012	
PNN		PCN	PAM	-223.012	450.023	175.690	11.438		0.749			
			IM	-209.081	428.161	173.686	12.0927	11.186	0.815	10.283	0.008	
PNN		PWC	PAM	-213.522	431.045	174.418	11.884		0.774			
			PAM	-206.503	417.006	177.102	10.396		0.774			
PNN		PWC2	IM	-205.865	421.729	175.494	4.217	8.984	0.698	34.424	5.743	
			IM	-226.754	463.508	179.383	2.624	10.951	0.741	19.181	0.928	

MOLECULAR ECOLOGY RESOURCES

		PAM	-236.056	476.113	179.704	11.724		0.764			
	PCN	PWC	PAM	-204.689	413.379	173.621	12.847		0.792		
			IM	-204.699	419.397	172.511	11.010	68.866	0.764	2.645	23.721
	PCN	PWC2	IM	-222.257	454.515	176.997	1.781	11.705	0.688	32.930	3.490
			PAM	-227.785	459.570	178.414	11.875		0.747		
	PCN	PSW	IM	-222.381	454.761	174.483	1.230	18.239	0.694	31.283	7.626
			PAM	-231.307	466.614	177.012	11.334		0.785		
	PWC	PWC2	IM	-214.582	439.162	173.315	10.381	16.970	0.837	0.433	9.814
			PAM	-218.713	441.427	174.934	11.831		0.792		
	PWC	PSW	IM	-213.585	437.169	174.067	10.468	99.796	0.883	0.298	9.708
			PAM	-216.763	437.526	176.123	11.409		0.814		
	PWC2	PSW	IM	-241.299	492.597	180.295	4.275	9.296	0.746	7.917	3.224
			PAM	-256.110	516.220	181.661	12.775		0.786		

Table S6: Variance Inflation Factor (VIF) values obtained for retained environmental variables per species: mean dissolved oxygen (DO_{mean}); mean Sea Surface Temperature (SST_{mean}), mean chlorophyll a concentration (CA_{mean}); mean Sea Surface Salinity (SS_{mean}).

Species	Environmental variable	VIF
<i>M. capensis</i>	DO _{mean}	1.007
	SST _{mean}	1.007
<i>M. paradoxus</i>	CA _{mean}	2.628
	DO _{mean}	1.457
	SS _{mean}	3.295

Table S7: Pearson correlation matrix obtained for *M. capensis* environmental variables: mean dissolved oxygen (DO_{mean}); mean Sea Surface Temperature (SST_{mean}).

	DO _{mean}	SST _{mean}
DO _{mean}	1	
SST _{mean}	-0.498	1

Table S8: Pearson correlation matrix obtained for *M. paradoxus* environmental variables: mean chlorophyll a concentration (CA_{mean}); mean dissolved oxygen (DO_{mean}); mean Sea Surface Salinity (SS_{mean}).

	CA _{mean}	DO _{mean}	SS _{mean}
CA _{mean}	1		
DO _{mean}	0.193	1	
SS _{mean}	-0.480	-0.194	1

MOLECULAR ECOLOGY RESOURCES

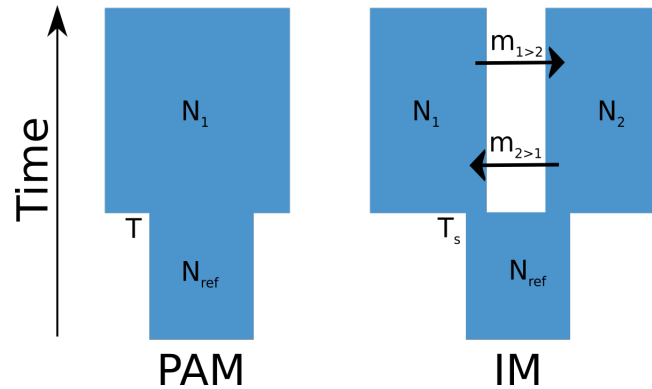


Figure S1: Schematic representation of the panmictic model (PAM) and isolation-with-migration model (IM). In PAM, an ancestral population of size N_{ref} goes through a change in population size (from N_{ref} to N_1) at a period of time T . In IM, an ancestral population of size N_{ref} split at a time T_s into two derived populations of size N_1 and N_2 , which are then connected by asymmetrical gene flow ($m_{1>2}$ and $m_{2>1}$).

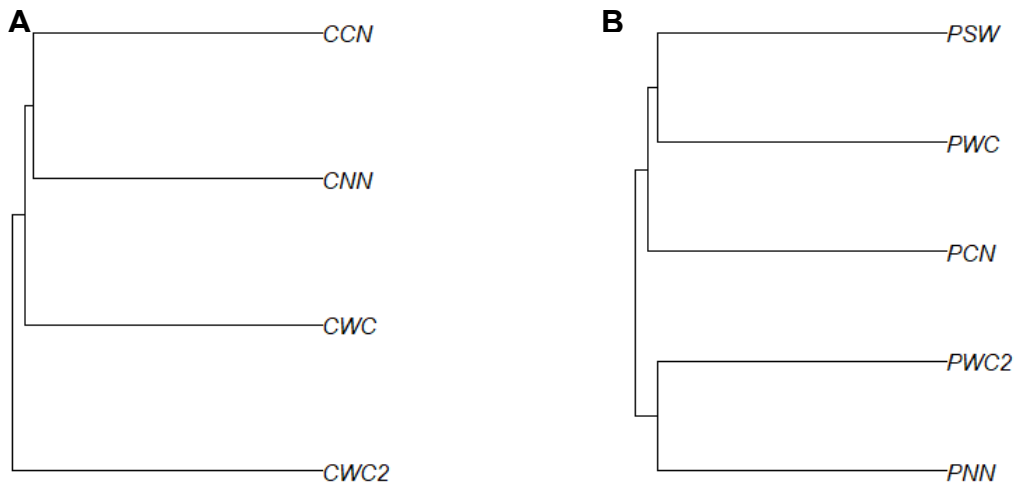


Figure S2: Hierarchical clustering trees for (A) *M. capensis* and (B) *M. paradoxus* based on the scaled covariance matrices Ω . Pool names as per Table S1.