Report on Hake Genetics

Stock assessment workshop

(2 December – 6 December 2013, Cape Town)

Compiled by: Conrad Matthee, Romina Henriques, Nina du Toit, and Sophie von der Heyden (Evolutionary Genomics Group, Stellenbosch University)

Collaborators: Marek Lipinski and Paulette Bloomer

Current state of Hake genetic stock structure:

In an attempt to define genetic stock structure in *M. paradoxus* and *M. capensis*, two research projects were initiated in two separate laboratories (at Stellenbosch and Pretoria Universities). The first involved mitochondrial DNA analyses performed at Stellenbosch; this study aimed to provide a first indication of population substructure in both species. Simultaneously, DNA of the same specimens used in the mtDNA analyses, were provided to the research group in Pretoria to perform microsatellite analyses on both species using species specific microsatellite primers designed by the laboratory in Pretoria. The latter approach is generally accepted to provide a more accurate/sensitive method to define differences in the distribution of genetic alleles. Both these research projects provided valuable information that can be used as a baseline for future investigations and a short summary of the outcome of each of these studies are provided below.

Published mitochondrial DNA stock structure results: The initial study utilized 644 hake sequences derived from the variable 5' region of the mtDNA control (311 *M. capensis* and 333 *M. paradoxus*). Samples were obtained between Lüderitz (southern Namibia) to south of Cape Point (South Africa). *Merluccius capensis* showed a high level of haplotypic diversity (107 haplotypes) while only 8 haplotypes were recovered for *M. paradoxus*. No evidence of mtDNA population structure could be found for *M. capensis*, however significant genetic differentiation was present in adult *M. paradoxus* (>3 years old). No significant population genetic structuring for detected for fish <3 years (~30cm). Two populations were identified which corresponded to Namibian and South African 'populations' (to the north and south of the Orange River). There was also significant structure along the South African coast. To establish the temporal stability of the *M. paradoxus* pattern a total of 1013 adult *M. paradoxus* (longer than 30cm), sampled over a period of three years were analysed (in addition to 134 adult fish taken from the 2005 study, 267 additional fish were included for the same period, 412 specimens were sampled in 2006 and 334 in 2007). For 2005, fish were from southern Namibia to the western Agulhas Bank, in 2006 we included the most northern distribution on the border with Angola, but none from southern Namibia. For 2007, samples ranged from southern Namibia to the western Agulhas Bank. Analyses of population structure revealed a similar northern (Namibian) southern (South African) differentiation but the boundaries were not as exact as in the first study. Notably, only two common haplotypes represent the

majority of the sampling effort (Fig 1).

In 2005, Namibia was significantly differentiated to SA1, SA2 and SA3 (Fst = 0.04-0.08, P < 0.05), but not SA4 (Fig 1). In 2006, the Cunene samples (no samples from southern Namibia were available), showed significant differentiation between SA3 and SA4 (Fst = 0.13-0.14, P < 0.03), but none of the other regions. In 2007, SA1 and SA3 were genetically differentiated. Combining the three single year data sets recovered significant structure between the Cunene River samples and all other sampling areas, except for southern Namibia (Fst = 0.13-0.19, P < 0.04), between southern Namibian samples and SA1, SA2 and SA4 (Fst = 0.02-0.04, P < 0.05) and between SA1 and SA3 (Fst = 0.009, P < 0.05). No significant inter-annual differentiation was present.

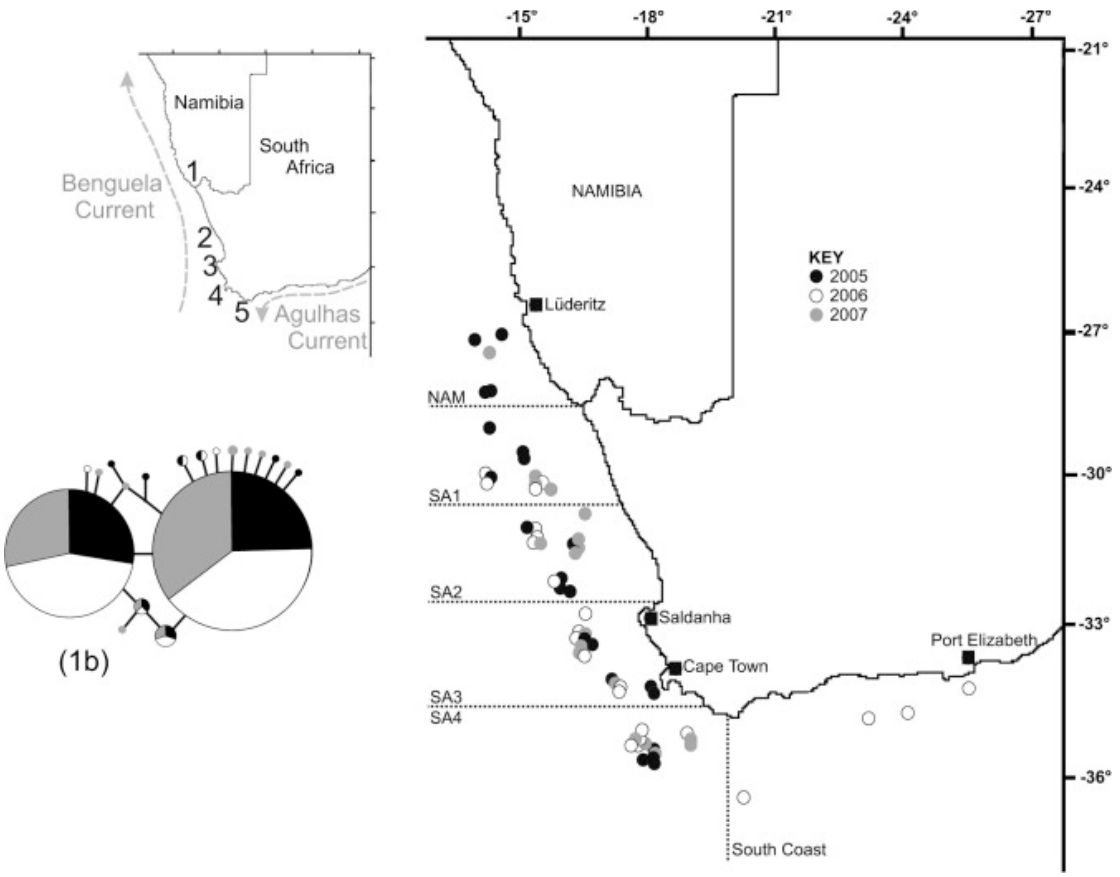

Figure 1: Figure taken directly from von der Heyden et al. 2010 showing sampling areas of *M. paradoxus*. The inset shows the two main southern African current systems. 1 = Orange River, the border between Namibia and South Africa, 2 = Hondeklip Bay, 3 = St. Helena Bay, 4 = Cape Point, 5 = Cape Agulhas. Main map shows areas sampled for the period 2005– 2007 (see key). Statistical parsimony haplotype network showing the two dominant haplotypes of 1013 *M. paradoxus* sampled. The size of the circles represents the relative frequency of individuals defining each haplotype; haplotypes are shaded according to year sampled.

Unpublished Microsatellite stock structure results: Since *M. paradoxus* showed mtDNA structure, and this species appears to be the older species within the southern African system, microsatellites were developed for *M. paradoxus*. Eight loci (six newly developed tetranucleotide loci and two European hake dinucleotide 3 loci) were genotyped in 209 M. paradoxus (57 trawls) and 222 M. capensis (24 trawls) adults collected during 2005 and 2006. Based on Bayesian STRUCTURE analyses the preliminary report indicated three likely stocks in *M. capensis* and one in *M. paradoxus*. The finding is in contrast to the mtDNA results. Small sample sizes per sampling trawl precluded a definitive outcome. Nevertheless, to determine regions where genetic patterns change abruptly, the software BARRIER was used. The geographic position of the significant genetic barriers in *M. capensis* was inferred to be (1) south of the Orange River on the west coast and (2) in the vicinity of Hermanus on the South African south coast. In general *M. paradoxus* showed higher levels of genetic variability than *M. capensis*, (14.75 alleles for *M. paradoxus* versus 11.13 for *M. capensis*) and higher levels of heterozygosity (average expected heterozygosity 0.72 versus 0.61). Deviations from Hardy-Weinberg expectations could be due to null-alleles or to recent demographic changes (as suggested by the mtDNA data).

Current Project: The present information given above offer a first step in providing genetic data useful for stock assessment in two hake species along the South African coastline. The data, however, suffer from several deficiencies and most of this relate to appropriate sampling and or the insufficient resolution of the mtDNA genetic markers used. To rectify the problem we aim to sample 20-40 stations (20 adult fish from each species covering the entire range of the species) over three consecutive seasons (March/April each year for three years). A mtDNA control region data set will be generated to test previous hypotheses that were inconclusive due to insufficient temporal/spatial sampling. For the microsatellites, it is recommended that at least 10 to 20 polymorphic loci is needed for the detection of stock differentiation. In addition should these loci not be informative enough, we will utilize new generation sequencing techniques to generate an additional 50+ polymorphic microsatellites useful for screening both hake species. The possibility also exist to extend this study and to utilize SNP data since newly developed markers are available for the European hake (Milano et al. 2013). The recent report of hybridization between the two Cape Hake species are also of concern and need to be further investigated Miralles et al. 2013). The completion date for the mtDNA and microsatellites analyses is the end of 2014 with a publication due by July 2015.

Progress made with the current Project:

A total of 897 *M. paradoxus* and 974 *M. capensis* were included for analyses (Supplementary Table 1). These samples were obtained during 2012 and 2013. For geographic analyses we defined 4 distinct geographic regions (Fig 2a,b). DNA extraction and sequencing has been performed on 1156 specimens and it is anticipated that this aspect will be completed by the time the 2014 samples arrived in the laboratory.

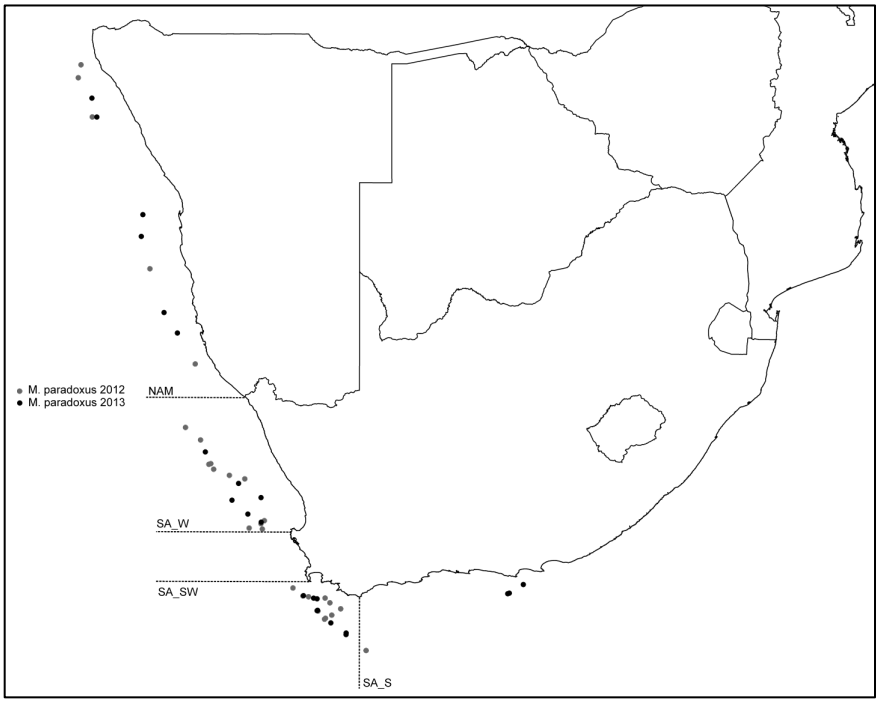

Figure 2a: Localities from which *M. paradoxus* was sampled during 2012 and 2013, indicating geographic region boundaries.

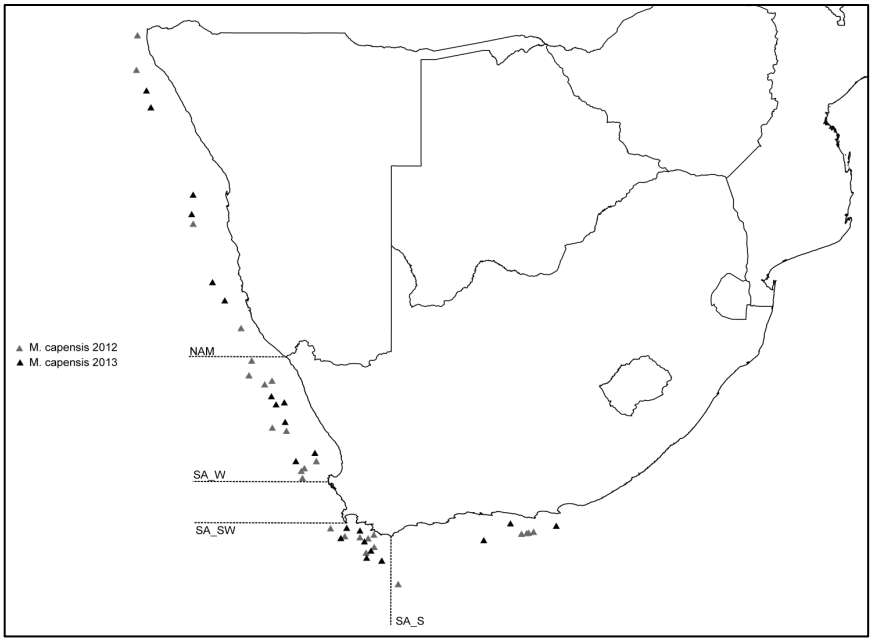

Figure 2b: Localities from which *M. capensis* was sampled during 2012 and 2013, indicating geographic region boundaries.

Preliminary mtDNA data analyses are based on 346 *M. paradoxus* and 364 *M. capensis* sampled mainly during 2012. Within *M. paradoxus,* 16 closely related haplotypes were detected (nucleotide diversity of 0.001; haplotypic diversity of 0.35). Similar to previous results, *M. capensis* showed a higher

level of diversity (nucleotide diversity of 0.003; haplotypic diversity of 0.69). and thus far, 53 haplotypes have been detected. In contrast to previous studies only showing significant differentiation in *M. paradoxus,* the preliminary analyses conducted for the 2012 / 2013 sample also show some level of significant population differentiation among the four sampling sites for *M. capensis* (Table 2). In light of the recent hybridization report it is important to also note that we have detected a similar anomaly during our recent analyses. A small percentage (<1%) of the sequenced individuals were considered "wrongly identified" but these may very well be signals of hybridization (the microsatellite analyses will reveal more information in this regard). Importantly, however, the identified *M. capensis* who carries mtDNA of *M. paradoxus* were mainly distributed in the SA south west sampling region (False Bay). At much lower frequency, we also detected mtDNA of *M. capensis* among two individuals identified as *M. paradoxus* (one from Namibia and one from False Bay). We have not confirmed these results but will do so before further analyses will be conducted. Since we did not regard hybridization as a possibility, these sequences were initially excluded from the mtDNA analyses presented below.

Table 2: Pairwise ΦST values indicting the differentiation among regions with *M. paradoxus* above and *M. capensis* below the diagonal. Significant values ($p < 0.05$) are indicated in bold.

	NAM	**SA_W**	**SA_SW**	**SA_S**
NAM	-	**0.028**	0.005	0.005
SA_W	0.003	-	0.014	**0.023**
SA_SW	0.001	**0.014**	-	0.003
SA_S	**0.025**	**0.057**	**0.03**	-

Eleven microsatellites, developed for the European Hake (*M. merluccius*) and for the Cape Hakes (*M. paradoxus* and *M. capensis*) have been screened for variability in both species. The entire set of microsatellites are amplifying in both hake species (Fig 3). Our preliminary diversity estimates are based on 10 *M. capensis* and 7 *M. paradoxus* specimens (Table 4).

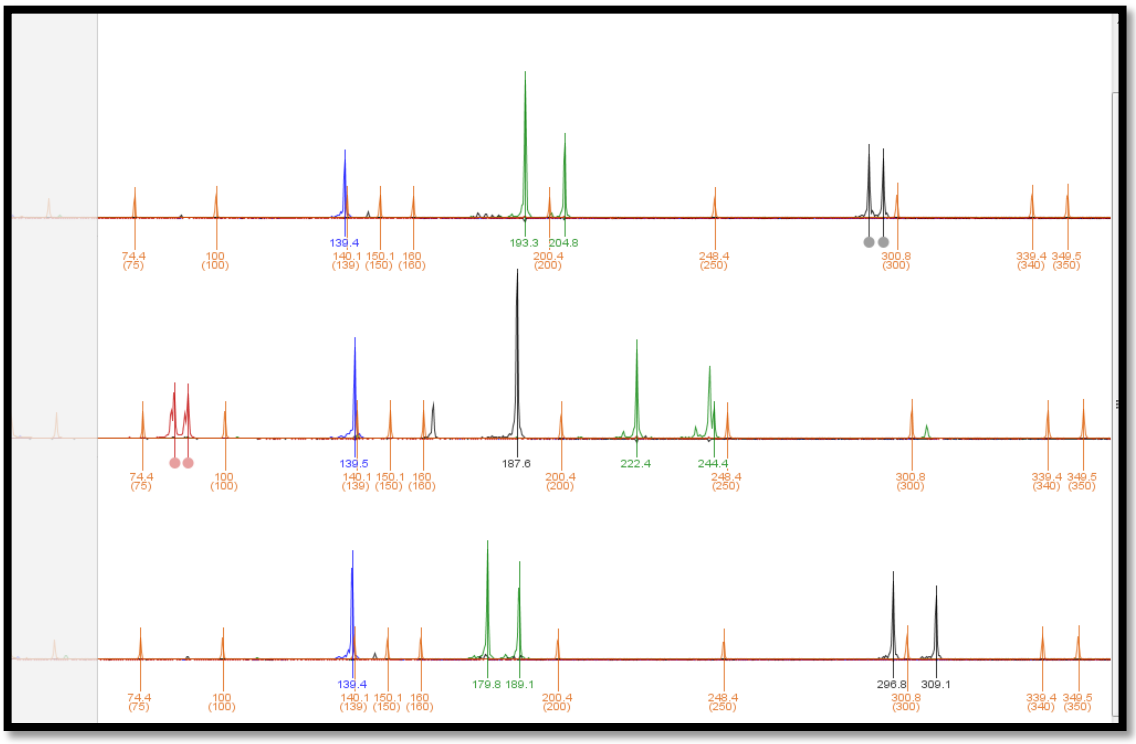

Figure 3: Example of a microsatellite chromatogram for hake indicating genotyping

Table 4: Microsatellite primers used, range (in bp) and number of alleles. Locus with an * were developed for the European hake.

Microsatellite loci	Allelic range	Number of Alleles	
		M. capensis	*M. paradoxus*
Mmer-hk3b*	300-350	2	4
Mmer-hk9b*	100-200	15	10
Mmer-hk20*	200-250	10	10
Mmer-hk29*	150-200	11	7
MP318	100-200	3	5
MP8450	150-300	12	10
MP51	100-200	4	2
MP374	70-100	2	3
MP8448	100-200	3	5
MP8748	150-250	7	8
MP8894	250-350	3	6

References:

Bloomer P, Dos Santos SMR, Oosthuizen CJ, Hoareau TB & Klopper AW 2009. “Development of microsatellite markers and screening of microsatellite locus variation in Cape hakes, *Merluccius paradoxus* and *M. capensis*, from the Namibian and South African coasts” Benefit Report

Milano I, Babbucci M, Cariani A, Atanassova M, Bekkevold D, Carvalho GR, Espinera M, Fiorentino F, Garofalo G, Geffen AJ, Nielsen EE, Ogden R, Patarnello T, Stagioni M, Tinti F, Bargelloni L (2013) Data from: Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius).* Dryad Digital Repository. doi:10.5061/dryad.7bn22

Miralles L, Machado-Schiaffino G., Garcia-Vazquez E. (in press). Genetic markers reveal a gradient of hybridization between cape hakes (*Merluccius capensis* and *Merluccius paradoxus*) in their sympatric geographic distribution. Journal of Sea Research.

von der Heyden S., Lipinski M.R. and Matthee C.A. 2007. Mitochondrial DNA analyses of cape hakes reveal an expanding, pan-mictic population for Merluccius capensis and population structuring for mature fish in Merluccius paradoxus. Molecular Phylogenetics and Evolution 42:517-527.

von der Heyden S., Lipinski M.R., Matthee C.A. 2007. Species-specific genetic markers for identification of early life history stages of Capehakes, Merluccius capensis and M. paradoxus in the southern Benguela Current. Journal of Fish Biology 70 (Supplement B):262-256.

von der Heyden, S., Lipinski, M., Matthee C.A. 2010. Remarkably low mtDNA control region diversity in an abundant demersal fish. Molecular Phylogenetics and Evolution 55: 1183-1188.

Supplementary Table 1: Localities from which *M. paradoxus* and *M. capensis* were sampled within each geographic region during 2012 and 2013, indicating those for which DNA has been extracted and the mtDNA control region has been sequenced.

Year	Area	Trawl	# *M. paradoxus*	# *M. capensis*	Latitude	Longitude	Extracted
2012	SA_SW	1		33	34.775000	19.403333	Y
2012	SA_SW	2		29	34.900000	19.210000	Y
2012	SA_SW	3	10		35.015000	19.083333	Y
2012	SA_SW	4		10	?	?	Y
2012	SA_SW	6	20		35.270000	18.711666	Y
2012	SA_SW	7	3		35.521666	18.916666	Y
2012	SA_SW	8	10		35.495000	18.946666	Y
2012	SA_SW	9	10	10	35.396666	19.141666	Y
2012	SA_SW	11	10	10	35.201666	19.416666	Y
2012	SA_SW	15	18	20	34.828333	18.423333	Y
2012	SA_SW	16	20	10	34.868333	18.933333	Y
2012	SA_SW	19	20		-34.80666	18.251666	Y
2012	SA_SW	25	12	12	34.556666	17.941666	Y
	SA_SW		133	134			
2012	SA_S	2		10	34.718333	24.565000	N
2012	SA_S	4		10	34.675000	24.798333	N
2012	SA_S	4		18	36.487000	20.224833	N
2012	SA_S	7		10	34.745000	24.391666	N
2012	SA_S	8	30		36.492000	20.203333	N
2012	SA_S	9		20	34.706666	24.635000	N
	SA_S		30	68			
2012	SA_W	54		20	32.225000	17.461666	Y
2012	SA_W	56	10	8	32.470000	17.056666	Y
2012	SA_W	57	10	9	32.811666	16.993333	Y
2012	SA_W	61	20		32.700000	16.585000	Y
2012	SA_W	64	10	10	32.565000	16.953333	Y
2012	SA_W	94		20	?	?	Y
2012	SA_W	96	10	10	31.176666	16.450000	Y
2012	SA_W	102	10	10	31.066666	15.973333	Y
2012	SA_W	113	1		30.706666	15.398333	Y
2012	SA_W	114	10		30.725000	15.338333	Y
2012	SA_W	115	10		30.876666	15.488333	Y
	SA_W		91	87			
2012	NAM (PN)	4		30	29.250000	15.183333	Y
2012	NAM (PN)	6	30		-29.583333	14.616666	Y
2012	NAM (PN)	25		30	27.616666	14.916666	Y

2012	NAM	25	30		27.616666	14.916666	N
2012	NAM	79	30		24.666666	13.516666	N
2012	NAM	89		30	24.000000	13.300000	N
2012	NAM	175		30	19.966666	11.866666	N
2012	NAM	176	30		19.966666	11.733333	N
2012	NAM	191	29		29.970000	15.080000	Y
2012	NAM	192		20	29.566666	15.715000	Y
2012	NAM	195		10	29.438333	15.963333	N
2012	NAM	195		30	17.466666	11.416666	N
2012	NAM	196		20	28.741666	15.275000	N
2012	NAM	205	30		18.350000	11.383333	N
2012	NAM	210		30	18.666666	11.383333	N
2012	NAM	211	30		18.750000	11.300000	N
	NAM		209	230			
2013	SA_S	1		15	34.966666	23.116666	Y
2013	SA_S	4	10		34.733333	24.583333	Y
2013	SA_S	8		20	34.466666	25.566660	Y
2013	SA_S	22	20		34.716666	24.633330	Y
2013	SA_S	27	20		34.450000	25.066660	Y
2013	SA_S	28		20	34.383330	24.016660	Y
	SA_S		50	55			
2013	SA_SW	1	15		35.259580	18.686330	N
2013	SA_SW	1		20	34.630000	18.935000	Y
2013	SA_SW	2		15	34.890000	18.288500	N
2013	SA_SW	2	10		34.890000	18.686330	Y
2013	SA_SW	2		10	34.761666	18.751666	Y
2013	SA_SW	3	20		35.637660	19.112660	Y
2013	SA_SW	3	10		34.870000	18.580000	Y
2013	SA_SW	9		20	35.575000	19.161660	N
2013	SA_SW	10		20	35.326600	19.308330	N
2013	SA_SW	12	20		36.000000	19.583330	N
2013	SA_SW	13	20		35.951666	19.585000	N
2013	SA_SW	14		10	35.678330	19.675000	N
2013	SA_SW	15		10	35.678333	19.666600	Y
2013	SA_SW	19	15		35.248333	18.701660	N
2013	SA_SW	22		15	35.008333	19.086660	Y
2013	SA_SW	26		10	34.536666	18.485000	Y
2013	SA_SW	29	20		34.795000	18.265000	Y
2013	SA_SW	85	10		31.755000	16.953330	Y
	SA_SW		140	130			
2013	SA_W	63	20		32.518333	16.958333	Y
2013	SA_W	75	20		32.266660	16.545000	Y
2013	SA_W	93	10		31.836660	16.053330	Y

2013	SA_W	100	20		31.316660	16.258330	Y
2013	SA_W	130	20		30.341660	15.230000	Y
2013	SA_W	76		20	32.225000	16.768333	Y
2013	SA_W	79		20	31.940000	17.413330	Y
2013	SA_W	104		10	30.870000	16.408333	Y
2013	SA_W	119		10	30.261666	16.100000	Y
2013	SA_W	120		20	30.190000	16.380000	Y
2013	SA_W	126		10	29.980000	15.941666	Y
	SA_W		90	90			
2013	NAM	42	17	30	26.658333	14.366666	N
2013	NAM	56	30	30	26.023330	13.953333	N
2013	NAM	92	17	30	23.666666	13.248333	Y
2013	NAM	109	30	30	22.991667	13.298333	N
2013	NAM	171	30	30	19.968333	11.873330	Y
2013	NAM	178	30	30	19.383330	11.721666	Y
	NAM		154	180			
Total sampled			897	974			
Total extracted			570	586			