Mar Biol
DOI 10.1007/s00227-009-1233-y

METHOD

Concordance of genetic and fin photo identification in the great white shark, *Carcharodon carcharias*, off Mossel Bay, South Africa

Chrysoula Gubili · Ryan Johnson · Enrico Gennari · W. Hermann Oosthuizen · Deon Kotze · Mike Meÿer · David W. Sims · Catherine S. Jones · Leslie Robert Noble

Received: 5 November 2008 / Accepted: 8 June 2009
© Springer-Verlag 2009

Abstract Visual identification of naturally acquired marks has been a popular if subjective method of animal identification and population estimation over the last 40 years. Molecular genetics has also independently developed objective individual marking techniques during the same period. Here, we assess the concordance of individual great white shark (*Carcharodon carharias*) dorsal fin recognition and identification, using seven microsatellite loci as the independent unbiased arbiter, over a period of 5 years. As a monitoring technique, fin photographs offer a very good individual identification key for white sharks over a relatively short period of time (5 years), matching with genetic data in about 85% of cases, whilst caution and a continuously updated database is required for animal recognition over a longer period.

Communicated by T. Reusch.

C. Gubili · C. S. Jones · L. R. Noble (✉)
Institute of Biological and Environmental Sciences,
University of Aberdeen, Zoology Building,
Tillydrone Avenue, Aberdeen AB24 2TZ, Scotland, UK
e-mail: l.r.noble@abdn.ac.uk

C. Gubili · D. W. Sims
The Laboratory, Marine Biological Association of the United Kingdom, Citadel Hill, Plymouth, Devon PL1 2PB, UK

R. Johnson
Department of Zoology and Entomology,
Mammal Research Institute, University of Pretoria,
Pretoria 0002, South Africa

R. Johnson · E. Gennari
Oceans Research, 92 Montague St, Mossel Bay 6500, South Africa

W. H. Oosthuizen · D. Kotze · M. Meÿer
Department of Environmental Affairs and Tourism,
Branch Marine and Coastal Management,
Private Bag 2, Roggebaai, South Africa

D. W. Sims
Marine Biology and Ecology Research Group,
School of Biological Sciences, University of Plymouth,
Drake Circus, Plymouth PL4 8AA, UK

Introduction

Photo identification is a technique mainly used on species that bear distinctive features, such as natural markings, which can be used to identify individuals. It has been used as a monitoring tool for a variety of marine and terrestrial species, although mostly applied to cetaceans (Karczmarski and Cockcroft 1998; Wilson et al. 1999; Calambokidis et al. 2004; Mizroch et al. 2004; Coakes et al. 2005), pinnipeds (Vincent et al. 2005), manatees (Langtimm et al. 2004), otters (Gilkinson et al. 2007), but also on cheetahs (Kelly 2001), and salamanders (Gamble et al. 2008). It is a relatively cheap, non-invasive technique allowing the physical "re-sampling" of an individual numerous times without artificially marking the studied species. This is vital for species that are difficult to tag because of their large size and intractability (Kohler and Turner 2001) in combination with weather conditions, or because they cannot retain the marks for the duration of the evaluation (Meekan et al. 2006; Gamble et al. 2008).

Since mid 1970s, when it was first used, photo identification has passed from film based photographs, with formation of slides and large photographic catalogues, to digitalisation and analysis of long-term databases using sophisticated image recognition software for faster more objective categorisation and individual recognition (Arzoumanian et al. 2005). The efficacy of individual

Published online: 24 June 2009

identification methods has escalated with technological advances. Digital photography is less laborious, more affordable and reliable, improving photographic quality (Markowitz et al. 2003), whilst recognition algorithms, although time consuming and costly to develop, avoid long-term commitment of financial and personnel resources; considered the most important drawbacks to photo identification (Hillman et al. 2003). Nevertheless, automation does not produce perfect results, since the final decision will be determined by the observer, introducing once again a degree of subjectivity (Araabi et al. 2000; Kelly 2001).

Photo identification can be used as part of a suite of recently implemented stress free and less invasive techniques for studying the population ecology and life history of species, particularly those under serious threat. It has been used for estimates of survival and population sizes (Wilson et al. 1999; Fujiwara and Caswell 2001; Langtimm et al. 2004; Mizroch et al. 2004; Castro and Rosa 2005; Bradshaw et al. 2007), sex or individual identification (Gowans et al. 2000; Kelly 2001; Hillman et al. 2003; Arzoumanian et al. 2005; Van Tienhoven et al. 2007), animal movement tracking and aggregations (Anderson and Goldman 1996; Sims et al. 2000; Vincent et al. 2005; Domeier and Nasby-Lucas 2007; Drouot-Dulau and Gannier 2007; Rowat et al. 2007) which can now be determined and monitored without the use of any physical tags or intervention, so common in the past (Kohler and Turner 2001).

The great white shark (*Carcharodon carcharias*, Linnaeus 1758) is elusive and capable of migrating long distances (Boustany et al. 2002; Bonfil et al. 2005), making photo identification techniques more attractive than conventional marking methods. Being economical and non-invasive, it allows the involvement of less experienced personnel (Castro and Rosa 2005). Nevertheless, despite its wide application to other marine species, particularly mammals, it has been used rarely with elasmobranchs (Sims et al. 2000). Over the last 15 years its use has focused primarily on species with very distinct characteristics that are classified as 'vulnerable' on the 2000 IUCN Red List for Threatened Species (Chapman et al. 2003). The white shark was one of the first elasmobranch species on which this approach was implemented (Anderson and Goldman 1996). The Australian Government successfully listed the white shark on Appendices I and II of the Convention on Migratory Species (CMS) in 2002. Efforts were orientated in classifying the legal status of the white sharks on Appendix II of the Convention on International Trade in Endangered Species of Fauna and Flora (CITES), a case that was finally materialised in 2004 after the cooperation between the Australian and Madagascan Governments. Klimley and Anderson (1996) were the first to verify the movements of white sharks in Californian waters using photographic evidence, reporting the existence of a database since 1988 for a 4 year period. More recently, Domeier and Nasby-Lucas (2007) managed to identify 78 individuals over a period of 5 years (2001–2005), allowing them to examine site fidelity at Guadalupe Island, Mexico. These authors used underwater photographs and video records of gill-slit flaps, pelvic and caudal fins, diverging from the established standard dorsal fin examination.

Despite the introduction and establishment of such methodologies, whilst there have been a couple of comparative studies in marine mammals (Gowans et al. 2000; Stevick et al. 2001), no studies testing the reliability of photographic identification using molecular techniques on elasmobranch skin-biopsies as the final arbiter have been conducted. Microsatellites, neutral codominant molecular genetic markers, have a range of applications in biology, medicine, forensics, molecular epidemiology, parasitology, population and conservation genetics, and genetic mapping (Chistiakov et al. 2006). Here, we assess the accuracy of photographic dorsal fin profiling by the objective approach of microsatellite genotyping each individual *C. carcharias* sampled in Mossel Bay, South Africa.

Materials and methods

Study site

Mossel Bay (34°11′S, 22°09′E) lies on the southern coast of South Africa. It is a well known white shark aggregation site. It houses a Cape fur seal (*Arctocephalus pusillus pusillus*) colony of over 4,000 pinnipeds. Mossel Bay's white shark distribution is characterised by three core areas: Seal Island, Hartenbos and Grootbrak.

Photo identification technique

Sharks were identified by sighting of dorsal fin and body marks when approaching the boat, drawn in by the smell of chum (mixture of fish products) and curiosity. No shark was rewarded intentionally with food for coming close to the boat for fear of conditioning the animal, despite the unlikely behavioural effect of feeding the sharks (Laroche et al. 2007). Sex, size, behaviour and individual unique body features (shape of a fin, notches, pigmentation, marks, bites, spots, scars, and deformities) of each shark were recorded during the encounter. Sizes were calculated by using the boat as an object of known length (7 m long) by an experienced observer, giving an approximation accurate to ±25 cm. Records of location and time were also taken while the shark's dorsal fin, its most diagnostic feature, was photographed. When possible, tissue and skin samples were collected using a biopsy needle. Samples were stored in 95%

ethanol at +4°C, or 20% DMSO saturated with NaCl at room temperature, until DNA extraction could be performed.

Photographs were taken using digital cameras with a minimum of six Megapixel resolution. Experienced and trained photographers were advised to take ID photographs when the subject was close to the boat, when the sea state (visibility) was good, and when the fin was perpendicular to the sea surface. Both sides of fin were taken when conditions permitted. Overall, 165 trips to anchorages within Mossel Bay were made from 10 May 2001 to 25 June 2005. During these trips, 527 sharks sightings were made, of which 314 sharks (60%) were successfully photographed.

All photographs were placed on a catalogue of all individuals, arranged by year. Each new photo was then compared with each other and those from all previous years. Individual shots were transposed over one another using Adobe Photoshop version 6.0 (Adobe Systems, San Jose, CA). In this way, it was possible to manipulate all photographs of an individual, including both the left and the right side of the dorsal fin. If a photograph was a positive match, it was added to those of a specific individual. If the match was negative, then the new animal was given an original number and entered into the catalogue. Confirmed matches were based on as many marks as possible to reduce the possibility of false positives. All images were first compared by an experienced observer, and then confirmed by a second.

Genetic identification

DNA extraction and microsatellite analysis

Total genomic DNA was extracted from muscle tissue and skin of all 110 samples using the standard phenol–chloroform extraction procedure (Sambrook et al. 1989). Polymerase chain reaction (PCR) amplifications were performed in a Biometra T-Gradient thermal cycler using primers for five South African great white shark microsatellite loci (Ccar1, Ccar6, Ccar9, Ccar13, and Ccar19) under the following conditions: reaction mix contained 10 ng template DNA, 1–2 mM $MgCl_2$ (Bioline, Inc), 200 µM of dNTPs (Bioline, Inc.), 1× NH_4 buffer (Bioline, Inc.), 0.3 µM of each primer, and 0.5 U *Taq* polymerase (Bioline, Inc.), denatured at 94°C for 5 min, followed by 36 cycles of 20 s at 94°C, 20 s at 60–62°C (depending on primers, see Pardini et al. 2000), and 30 s at 72°C, and a final extension of 10 min at 72°C. A new species specific primer (Ccar6.27x) was also developed and used in the analysis (F 5′-GAGCATGTGTGGGAGCGAAAG-3′ and R 5′-TGGGACGATTCTGCCATTCTCTC-3′) under the following conditions: the reaction mix contained 10 ng template DNA, 1.5 mM $MgCl_2$ (Bioline, Inc.), 200 µM of dNTPs (Bioline, Inc.), 1× NH_4 buffer (Bioline, Inc), 0.3 µM of each primer, and 0.5 U *Taq* polymerase (Bioline, Inc.), denatured at 94°C for 5 min, followed by 36 cycles of 30 s at 94°C, 30 s at 54°C, and 30 s at 72°C, with a final extension of 10 min at 72°C. Finally, a non species specific primer (Iox-10) from the shortfin mako, *Isurus oxyrinchus*, was also used as described in Schrey and Heist (2002). All PCR conditions and microsatellite characteristics are presented in Table 1. All PCR products were genotyped on a LICOR DNA sequencer 4200L (Global Edition IR^2) and all runs were downloaded to the e-Seq V2.0 software program. Fragments and their sizes were determined against positive controls (other South African samples) and a molecular weight ladder (HyperLadder IV, 5 µl, Bioline).

Concordance of methods

Photographic identification was followed by genetic identification, which was conducted as a double blind test to minimise bias. There had been no exchange of results prior to this analysis. Over a period of 5 years, 314 sharks were photographed and 110 biopsies collected.

Table 1 Genetic variation measures of seven microsatellite markers for *C. carcharias* sampled in Mossel Bay, South Africa (Ccar1, Ccar6, Ccar9, Ccar13, and Ccar19 are from Pardini et al. (2000); Iox-10 is from Schrey and Heist (2002))

Mossel Bay samples										
Locus	Accession No	*k*	Ho	He	HWE	F_{is}	$P_{(ID)unbiased}$	$P_{(ID)sibs}$	$P_{(ID)unbiased}$#	$P_{(ID)sibs}$#
Ccar1	AF216865	6	0.682	0.709	ns	0.038	1.287×10^{-1}	4.308×10^{-1}	1.287×10^{-1}	4.308×10^{-1}
Ccar9	AF216866	15	0.909	0.869	ns	−0.046	2.608×10^{-2}	3.248×10^{-1}	3.355×10^{-3}	1.399×10^{-1}
Ccar13	AF184087	12	0.685	0.778	ns	0.120	7.303×10^{-2}	3.822×10^{-1}	2.450×10^{-4}	5.348×10^{-2}
Ccar19	AF216864	3	0.527	0.524	ns	−0.006	3.403×10^{-1}	5.749×10^{-1}	8.337×10^{-5}	3.075×10^{-2}
Iox-10	AF426735	5	0.700	0.755	ns	0.073	9.774×10^{-2}	3.999×10^{-1}	8.148×10^{-6}	1.230×10^{-2}
Ccar6.27x		3	0.528	0.457	ns	−0.155	3.329×10^{-1}	6.076×10^{-1}	2.713×10^{-6}	7.472×10^{-3}
Ccar6	AF184085	4	0.633	0.628	*	−0.008*	2.035×10^{-1}	4.896×10^{-1}	5.520×10^{-7}	3.659×10^{-3}

k number of alleles, *Ho* observed heterozygosity, *He* expected heterozygosity, F_{is} inbreeding coefficient, deviations from HWE, $P_{(ID)}$ probability of identity, * $p < 0.05$, *ns* not significantly different from 0, # cumulative values for $P_{(ID)}$

Genotyping errors, such as allelic dropout, stutter bands, and null alleles, were tested for all loci using MICRO-CHECKER 2.2.3 (van Oosterhout et al. 2004). Tests for significant deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were performed using GENEPOP 4.0.7 (Rousset 2007). Inbreeding coefficients (F_{is}) were estimated to evaluate levels of genetic diversity, and number of alleles (N_A), as well as observed (H_O), and expected heterozygosities (H_E) (Nei 1987) were calculated using CERVUS 3.0 (Kalinowski et al. 2007).

Firstly, the rate of genotyping error was assessed to identify any bias in the results. Bonin et al. (2004) stressed that all genetic data sets include some erroneous genotypes due to various causes. All data were double-checked to eliminate scoring errors. The Mossel Bay dataset included 110 individuals that had been scored for seven loci. Sixteen samples (14.55%) were re-genotyped for all loci, and a comparison made between old and re-scored multilocus profiles to estimate the genotype error rates as the number of allelic mismatches.

CERVUS 3.0 (Kalinowski et al. 2007) was used for identification based on genetic analysis of all sampled *C. carcharodon* to investigate and highlight all the exact and close matches in the dataset. Cervus 3.0 is a maximum likelihood approach commonly used in parental assignment and is a useful check for identifying repeat sampling. During the identity check, Cervus reads the genotypes from an individual and compares them against every genotype in the data file. Cervus records identifications, and repeated genotypes, offering a "fuzzy" matching function (not 100% matches), as well as perfect matches; it is also capable of including sex in the search. This method of assignment is based on allele frequencies generated at the beginning of the analysis for the Mossel Bay white shark population (overall 110 individuals, based on previous and new samples, excluding animals that have been biopsied repeatedly). All mismatches were examined by eye to confirm that allelic drop out did not contribute to the false exclusion of individuals.

The probability of individual identity, $P_{(ID)}$, for all seven loci was calculated according to Waits et al. (2001) with GIMLET (Valière 2002); as the probability of two individuals/animals that belong to one population, and drawn randomly sharing the same genotypic profile at multiple loci (Waits et al. 2001). GIMLET allows the calculation of both the upper $P_{(ID)sib}$ (is the $P_{(ID)}$ of a population where siblings are found and included; Evett and Weir 1998), and lower $P_{(ID)unbiased}$ (is the $P_{(ID)}$ after sample size corrections; Paetkau et al. 1998) boundaries of $P_{(ID)}$, assuming that the population studied does not deviate from Hardy-Weinberg expectations. Waits et al. (2001) showed that the observed $P_{(ID)}$ is found between the $P_{(ID)sib}$, and $P_{(ID)unbiased}$, and is estimated by computing the proportion of all possible pairs of individuals that have identical genotypes.

Results

Only 128 photographs were suitable for further investigation, since these photographs were of biopsied individuals. In total, 84 individuals were recognised, of these there have been 28 repeated identifications of the same animals, and 56 of new individuals. Two individuals were identified on up to five different occasions throughout 5 years of sampling. The longest individual identification was separated by 1,596 days, and the shortest by only 1 day.

One locus, Ccar6 was found to deviate from HWE. However, little in the way of allelic dropout, stutter bands, or null alleles were reported at this locus by MICRO-CHECKER. However, there was a possibility of null alleles at locus Ccar13, and LD between this locus and Ccar19 was identified. Nevertheless, on a different and not so geographically restricted dataset, deviations from HWE and LD were not detected (Pardini et al. 2001), suggesting the presence of related individuals in Mossel Bay. There were 3–15 alleles per locus, while observed and expected heterozygosities ranged from 0.527 to 0.909 and 0.457 to 0.869, respectively, and F_{is} values from −0.155 to 0.073 (Table 1). Ignoring the failed amplification in the second genotyping, two genotyping errors at different loci were identified, resulting in a 0.01089 error rate.

Association of photo with genetic identification was made on an available sub sample (15 animals repeatedly photographed and biopsied) of 28 individuals identified by separate photographic matches. Analysis using Cervus 3.0, showed the existence of 25 perfect identifications of biopsied animals, revealing 11 individuals out of total 37 photographs that have been sampled more than once from between 2001 and 2005 (Table 2). Each identification was checked further by eye. On four occasions there had been discrepancies in concordance of the two techniques (Table 2).

These four occasions on which discrepancies were found, were false positives (two sightings of different animals considered as the same individual). An additional analysis using the 'fuzzy matching' procedure of Cervus was run to determine if false negative errors (two sightings of the same animal considered to be different individuals) were evident from biopsies sampled on two different occasions. Identity analysis of biopsied animals revealed 22 cases of sharks with five, and three cases with six matching loci, respectively, proving the samples from each sampling date were in fact of different individuals, suggesting there were no records of false negatives in our dataset.

Table 2 Concordance of photo identification and genotyping for 15 possible individual *C. carcharias* from 37 biopsied and photographed samples from Mossel Bay, following "identity analysis" and match in Cervus

Genetic ID	Photo ID	Date of capture	Name	Sex	Ccar1	Ccar6	Ccar9	Ccar13	Ccar19	Iox-10	Ccar6.27x
GWSA 03/05	1	15/06/2002	Inverse nipple	Male	169–159	219–219	234–214	272–272	212–210	125–121	170–170
GWSA 03/08	11	16/06/2002	Inverse nipple	Male	169–159	219–219	234–214	272–272	212–210	125–121	170–170
GWSA 03/01	2	17/06/2002	Nipple step	Female	169–159	219–219	236–220	292–272	212–210	129–127	174–170
GWSA 03/21	1	18/08/2002	Nipple step	Female	169–159	219–219	236–220	292–272	212–210	129–127	174–170
GWSA 03/12	5	20/08/2002	Nipple step	Female	169–159	219–219	236–220	292–272	212–210	129–127	174–170
GWSA 04/94	6	26/02/2004	Nipple step	Female	159–159	221–219	220–216	292–292	212–210	127–125	170–170
GWSA 03/47	1	23/06/2001	Monster Rosie	Female	169–161	225–219	220–214	290–290	212–212	125–125	174–170
AU05-0498	6	7/11/2005	Monster Rosie	Female	169 161	225–219	220–214	290–290	212–212	125–125	174–170
GWSA 03/74	5	24/05/2003	Speedy	Female	167 159	225–221	224–214	292–272	212–210	129–127	170–170
GWSA 04/88	1	24/02/2004	Speedy	Female	171 159	225–221	220–220	290–290	212–210	125–125	170–170
GWSA 03/48	16	23/06/2001	Top notch	Female	169–159	219–219	222–220	292–272	212–210	129–121	178–170
GWSA 03/18	1	26/08/2002	Top notch	Female	169–159	219–219	222–220	292–272	212–210	129–121	178–170
AU05-0472	11	22/06/2004	Top notch	Female	169–159	219–219	222–220	292–272	212–210	129–121	178–170
GWSA 03/46	2	23/06/2001	Half moon	Female	161–159	225–219	236–228	292–290	212–210	125–125	170–170
AU05-0455	18	2/9/2004	Half moon	Female	161–159	225–219	236–228	292–290	212–210	125–125	170–170
GWSA 03/16	16	20/08/2002	Square notch	Male	167–159	225–221	224–214	292–272	212–210	131–129	170–170
GWSA 03/73	002a	31/05/2003	Square notch	Male	171–159	225–221	220–220	292–290	212–210	127–127	170–170
GWSA 03/22	5	10/11/2003	Mystery fin	Female	169–167	225–219	220–214	292–272	212–210	127–127	178–170
AU05-0453	13	2/9/2004	Mystery fin	Female	169–167	225–219	220–214	292–272	212–210	127–127	178–170
AU05-0479	1	28/09/2005	Mystery fin	Female	169–167	225–219	220–214	292–272	212–210	127–127	178–170
AU05-0486	2	18/10/2005	Mystery fin	Female	169–167	225–219	220–214	292–272	212–210	127–127	178–170
GWSA 03/40	012a	26/06/2001	Glasses	Female	171–171	225–225	220–216	286–278	212–210	131–121	178–174
GWSA 04/89	3	24/02/2004	Glasses	Female	171–171	225–225	220–216	286–278	212–210	131–121	178–174
AU05-0462	4	7/6/2005	Glasses	Female	171–171	225–225	220–216	286–278	212–210	131–121	178–174
AU05-0468	12	21/06/2004	Round top	Female	169–161	225–219	236–220	292–290	212–212	127–127	174–170
AU05-0464	8	10/6/2005	Round top	Female	169–161	225–219	236–220	292–290	212–212	127–127	174–170
GWSA 03/07	3	15/06/2002	Top notch steps	Female	159–159	225–221	238–220	278–272	212–210	127–127	178–174
GWSA 03/15	11	22/08/2002	Top notch steps	Female	159–159	225–221	238–220	278–272	212–210	127–127	178–174
GWSA 04/93	4	26/02/2004	Santa	Female	169–159	219–219	236–220	292–272	212–210	129–127	174–170
AU05-0497	5	7/11/2005	Santa	Female	159–159	221–219	220–216	292–292	212–210	127–127	170–170
AU05-0470	2	22/06/2004	Big rosie WTT	Female	169–161	225–221	220–214	292–290	212–210	127–125	170–170
AU05-0485	2	6/10/2005	Big rosie WTT	Female	169–161	225–221	220–214	292–290	212–210	127–125	170–170
AU05-0475	5	23/06/2004	Gill scar	Female	171–169	225–219	232–212	292–292	212–210	129–127	170–170
AU05-0450	3	2/9/2004	Gill scar	Female	171–169	225–219	232–212	292–292	212–210	129–127	170–170
AU05-0457	2	12/5/2005	Gill scar	Female	171–169	225–219	232–212	292–292	212–210	129–127	170–170
GWSA 03/04	2	18/06/2002	Tadpole rosie	Female	161–159	219–219	214–212	292–272	218–210	129–127	170–170
GWSA 04/75	3	15/05/2004	Tadpole rosie	Female	161–159	219–219	214–212	292–272	218–210	129–127	170–170

Animals in bold are those that lack concordance between genetic and photographic identification. Loci are characterised in Table 1

The $P_{(ID)unbiased}$ was calculated originally for all seven loci as 5.52×10^{-7}, and the probability of individual identity among siblings as 3.66×10^{-3} (Table 1). The outcome of estimating the upper and lower limits of $P_{(ID)}$ changed when loci Ccar6 was excluded (which did not conform to HWE). $P_{(ID)unbiased}$ was 2.71×10^{-6}, and $P_{(ID)sibs}$ was 7.47×10^{-3} (Table 1), therefore two individuals have a less than 1% chance of sharing the same multilocus profile.

All individuals exhibiting lack of concordance between genetic and photographic methodologies were rechecked by eye to confirm the accuracy of genotyping. After thorough examination by both observers of the first and consecutive photographic sampling, the photographs 005-20082002 and 006-26022004 belonged to "Nipple step", 005-24052003 and 001-24022004 was "Speedy", 016-20082002 and 002a-31052003 came from "Square notch", and 004-26022004 and 005-07112005 was "Santa" (Fig. 1).

Fig. 1 Photographic verification records at the first and second photo and biopsy sampling, following the lack of concordance with genetic analysis: "Speedy", "Santa", "Nipple Step", and "Square Notch" identification records, respectively, 005-24052003 (**a**) and 001-24022004 (**b**); 004-26022004 (**c**) and 00-07112005 (**d**); 005-20082002 (**e**) and 006-26022004 (**f**); 016-20082002 (**g**) and 002a-31052003 (**h**)

Discussion

Overall, comparison of genetic and photographic identification showed the approaches to be generally concordant. The application of microsatellite markers is an effective and independent unbiased arbiter for assessing veracity of persistent natural markings for the identification of individual animals. Genotype identification mistakes are minimised by the use of a standard size ladder run alongside the genotype of the studied individual, whilst the probability of two individuals having the same genotypes at these seven loci is less than 1% (Table 1), making false inclusion highly

unlikely, even when siblings are present in a population. Nevertheless, Anderson et al. (2002) (reporting an individual probability of identity of 8.189×10^{-17} for humpback whales) were able to identify one individual sampled twice using the same number of microsatellites, despite the application of the Paetkau and Strobeck (1994) biased formula (which does not incorporate a sample size correction). However, photographic identification remains a subjective method, since the final decision on identification is taken by the observer based on experience and expertise.

Discrepancies due to mismatches among individuals considered to be the same animal can be detected with genotype scoring. The resolution in mismatch recognition is also improved by the degree of polymorphism at loci used in the analysis. The more polymorphic loci used, the smaller the probability of two individuals sharing a specific allele. The number of alleles varied between the seven loci from low (Ccar19 and Ccar6.27x had 3 alleles each) to high (15 and 12 alleles for Ccar9 and Ccar13, respectively). In this study, genotyping and photographic mismatches were apparent using both highly polymorphic and less polymorphic loci. In all cases, discrepancies were detected for at least four loci per pairwise multilocus profile comparison (Table 2), justifying the identification of individuals, even with an error rate of 0.01089.

Six pairwise photographic discrepancies out of 37 cases (16.22%) were noted, whereas comparison of genetic and photographic identification of 15 possible individuals showed four false positives, which translates into a 26.67% of the sample used in this study, or 74.33% of individual successful matches. These findings were similar to those of Stevick et al. (2001) for humpback whales, where he discovered 88 discrepancies between photographic and genetic identifications over 414 cases (21.26%) prior to correction. However, genetic analysis demonstrated an extremely low rate (3.3%) of animals miscategorised by sex, in northern bottlenose whales, *Hyperoodon ampullatus*, suggesting the technique is applicable for sex identification (Gowans et al. 2000). Nevertheless, the latter study was based on a species trait, melon morphology, which is indicative of sexual maturity and not on naturally acquired marks. A more reliable dorsal fin identification of white sharks may be obtained with a larger sample size, but caution is recommended for smaller datasets (Gowans and Whitehead 2001).

The application of photo identification is extremely valuable for the estimation of abundance, particularly in areas where the population size is unknown but is considered to be relatively small. The existence of errors has been shown to inflate abundance or bias survival estimates (Langtimm et al. 2004; Meekan et al. 2006), and Stevick et al. (2001) suggest bias can be minimised substantially with better quality photographs. Quality is one factor biasing animal identification, and has been reported previously in cetacean studies, motivating researchers to evaluate the accuracy of photo ID more objectively. When genetic tools cannot be implemented as a calibration, computer programs can be applied successfully in certain species. Arzoumanian et al. (2005) and Van Tienhoven et al. (2007) developed the first computer based software for shark identification with success rates of 90% and 95%, respectively. Both studies utilised the existence of discrete natural marks and patterns (whale and spotted raggedtooth sharks); unfortunately, white sharks bear no such patterns. "Finscan", developed for identifying marine mammals by comparing dorsal fin pictures with up to 75% correct identification (within the first three or four suggested matches), was used on white sharks (Hillman et al. 2003). However, this was a preliminary analysis utilising a small database, a larger sample size might have altered the success rate.

Image quality is a major factor affecting the efficacy of the analysis. Many studies have shown errors due to poor quality photographs (Gowans and Whitehead 2001), the authors stressing the importance of photographic conditions and the experience of the observer using both computer-aided programs and naked eye identification (Kelly 2001; Stevick et al. 2001; Arzoumanian et al. 2005; Gilkinson et al. 2007; Van Tienhoven et al. 2007; Gamble et al. 2008). Reduced image clarity due to water turbidity, bad weather conditions, photographic angles, body position, and comparison of both sides of the fin may explain unsuccessful recognition of individual sharks in the Van Tienhoven et al. (2007) study at least. In our case, all images were clear with a good photographic angle, except those of "Speedy" (Fig. 1), which were still clear enough to positively identify the same individual.

The process of classification is labour intensive, but could still lead to misidentification of animals (Kelly 2001; Hillman et al. 2003). Genetic and photo mismatches might occur through accidental mislabelling of photographs or sample tubes, during sampling or data analysis. Fatigue handling a dataset is correlated to the number of photographs used (Stevick et al. 2001). Human error has always been one of the major sources of inconsistency, so discrepancies prompted both genetic and photographic scoring to be re-examined.

False positives can also be attributed to the comparison of both sides of an individual. Domeier and Nasby-Lucas (2007) recognise the necessity of photographing both sides of an animal in order to catalogue it, because of differences between pigment patterns in all three regions they examined (gill flaps, pelvic and caudal fins) on the right and left sides. Another characteristic of sharks is their capacity to heal quickly (Fitzpatrick et al. 2006; Domeier and Nasby-Lucas 2007). Even large wounds can become indistinguishable, which is why pigment patterns are more reliable for individ-

ual identification. Nevertheless, during the five years of documentation and catalogue creation not only wounds, but also natural marks and even pigmentation were seen to alter (Domeier and Nasby-Lucas 2007), leading to false positives. Nevertheless, mis-matches can also be caused by new physical scarring and deformities (Castro and Rosa 2005). Dufault and Whitehead (1995) reported that difficulty of re-identifying an individual using natural markings increases with time, with major changes observed in sperm whales. This might also be the case for white sharks, hence, a constantly updated photographic database of animals will improve accuracy and reliability, allowing greater temporal stability in individual identification. Changeability over time and unreliability of natural marks strongly suggest application of a single feature for individual identification is unlikely to be useful, particularly in long-term recognition studies (Karczmarski and Cockcroft 1998).

Carcharodon carcharias, like all "charismatic megafauna", are usually tagged with any "pioneering" non-invasive method suggested to leave their behaviour or relationship with other individuals unaltered, and to not injure the animal (Markowitz et al. 2003). However, to test the reliability of such approaches a similarly non-invasive but objective technique is required. The main focus of this study was to evaluate the accuracy and utility of photographic identification of individual white sharks, a species with no distinctive patterns, based on naturally acquired marks. The outcome of the photographic analysis was not used to make estimates of abundance, or for any other potential applications already mentioned, as these are the subjects of other studies. Our findings emphasise that attention to photographic quality and site of character(s) are needed when analysing large datasets of sharks, and that previous photo identification work should be interpreted cautiously given the rate of mismatches detected if photographs are used in isolation. Ultimately, efforts incorporating photographic records of individuals with the corresponding genetic profiling would produce a more reliable means to assess the strategies and population dynamics of *C. carcharias*.

Acknowledgments This work was supported by Aberdeen University and the Marine Biological Association. We are grateful to two referees for valuable comments and Linda Sönderberg for her help in the laboratory. This work complies with the current laws of the United Kingdom.

References

Anderson SD, Goldman KJ (1996) Photographic evidence of white shark movements in California waters. Calif Fish Game 82(4):182–186

Anderson MJ, Hinten G, Paton D, Baverstock PR (2002) A model for the integration of microsatellite genotyping with photographic identification of humpback whales. Mem Qld Mus 47(2):451–457

Araabi BN, Kehtarnavaz N, McKinney T, Hillman G, Würsig B (2000) A string matching computer-assisted system for dolphin photo identification. Ann Biomed Eng 28:1269–1279

Arzoumanian Z, Holmberg J, Norman B (2005) An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*. J Appl Ecol 42:999–1011. doi:10.1111/j.1365-2664.2005.01117.x

Bonfil R, Meÿer M, Scholl MC, Johnson R, O'Brien S, Oosthuizen H, Swanson S, Kotze D, Paterson M (2005) Transoceanic migration, spatial dynamics, and population linkages of white sharks. Science 310:100–103. doi:10.1126/science.1114898

Bonin A, Bellemain E, Bronken Eidesen O, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261–3273

Boustany AM, Davis SF, Pyle P, Anderson SD, Le Boeuf BJ, Block BA (2002) Expanded niche for white sharks. Nature 415:35–36. doi:10.1038/415035b

Bradshaw CJA, Mollet HF, Meekan MG (2007) Inferring population trends for the world's largest fish from mark-recapture estimates of survival. J Anim Ecol 76:480–489. doi:10.1111/j.1365-2656.2006.01201.x

Calambokidis J, Steiger GH, Ellifrit DK, Troutman B, Bowlby CE (2004) Distribution and abundance of humpback whales (*Megaptera novaeangliae*) and other marine mammals off the northern Washington coast. Fish Bull 102:563–580

Castro ALF, Rosa RS (2005) Use of natural marks on population estimates of the nurse shark *Ginglymostoma cirratum*, at the Atol das Rocas Biological Reserve, Brazil. Environ Biol Fish 72:213–221. doi:10.1007/s10641-004-1479-7

Chapman DD, Abercrombie DL, Douady CJ, Pikitch EK, Stanhope MJ, Shivji MS (2003) A streamlined, bi-organelle, multiplex PCR approach to species identification: application to global conservation and trade monitoring of the great white shark, *Carcharodon carcharias*. Conserv Genet 4:415–425

Chistiakov DA, Hellemans B, Volckaert FAM (2006) Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. Aquaculture 255:1–29. doi:10.1016/j.aquaculture.2005.11.031

Coakes A, Gowans S, Simard P, Giard J, Vashro C, Sears R (2005) Photographic identification of fin whales (*Balaenoptera physalus*) off the Atlantic coast of Nova Scotia, Canada. Mar Mamm Sci 21(2):323–326

Domeier ML, Nasby-Lucas N (2007) Annual re-sightings of photographically identified white sharks (*Carcharodon carcharias*) at an eastern Pacific aggregation site (Guadalupe Island, Mexico). Mar Biol 150:977–984. doi:10.1007/s00227-006-0380-7

Drouot-Dulau V, Gannier A (2007) Movements of sperm whale in the western Mediterranean Sea: preliminary photo-identification results. J Mar Biol Ass UK 87:195–200. doi:10.1017/S0025315407054860

Dufault S, Whitehead H (1995) An assessment of changes with time in the marking patterns used for photoidentification of individual sperm whales, *Physeter macrocephalus*. Mar Mamm Sci 11(3):335–343

Evett IW, Weir BS (1998) Interpreting DNA evidence: statistical genetics for forensic scientists. Sinauer, Sunderland

Fitzpatrick B, Meekan M, Richards A (2006) Shark attacks on a whale shark (*Rhincodon typus*) at Ningaloo reef, Western Australia. Bull Mar Sci 78(2):397–402

Fujiwara M, Caswell H (2001) Demography of the endangered North Atlantic right whale. Nature 414:537–541

Gamble L, Ravela S, McGarigal K (2008) Multi-scale features for identifying individuals in large biological databases: an application of pattern recognition technology to the marbled salamander *Ambystoma opacum*. J Appl Ecol 45:170–180. doi:10.1111/j.1365-2664.2007.01368.x

Gilkinson AK, Pearson HC, Weltz F, Davis RW (2007) Photo-identification of sea otters using nose scars. J Wildlife Manage 71(6):2045–2051. doi:10.2193/2006-410

Gowans S, Whitehead H (2001) Photographic identification of northern bottlenose whales (*Hyperoodon ampullatus*): sources of heterogeneity from natural marks. Mar Mamm Sci 17(1):76–93

Gowans S, Dalebout ML, Hooker SK, Whitehead H (2000) Reliability of photographic and molecular techniques for sexing northern bottlenose whales (*Hyperoodon ampullatus*). Can J Zool 78:1224–1229

Hillman GR, Würsig B, Gailey GA, Kehtarnavaz N, Drobyshevsky A, Araabi BN, Tagare HD, Weller DW (2003) Computer-assisted photo-identification of individual marine vertebrates: a multi-species system. Aquat Mamm 1:117–123

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106. doi:10.1111/j.1365-294X.2007.03089.x

Karczmarski L, Cockcroft VG (1998) Matrix photo-identification technique applied in studies of free-ranging bottlenose and humpback dolphins. Aquat Mamm 24.3:143–147

Kelly MJ (2001) Computer-aided photograph matching in studies using individual identification: and example from Serengeti cheetahs. J Mammal 82(2):440–449

Klimley AP, Anderson SD (1996) Residency patterns of white sharks at the South Farallon Islands, California. In: Klimley AP, Ainley DG (eds) Great white sharks: the biology of *Carcharodon carcharias*. Academic, San Diego, pp 365–373

Kohler NE, Turner PA (2001) Shark tagging: a review of conventional methods and studies. Environ Biol Fish 60:191–223

Langtimm CA, Beck CA, Edwards HH, Fick-Child KJ, Ackerman BB, Barton SL, Hartley WC (2004) Survival estimates for Florida manatees from the photo-identification of individuals. Mar Mamm Sci 20(3):438–463

Laroche RK, Kock AA, Dill LM, Oosthuizen WH (2007) Effects of provisioning ecotourism activity on the behaviour of white sharks *Carcharodon carcharias*. Mar Ecol Prog Ser 338:199–209

Markowitz TM, Harlin AD, Würsig B (2003) Digital photography improves efficiency of individual dolphin identification. Mar Mamm Sci 19(1):217–223

Meekan MG, Bradshaw CJA, Press M, McLean C, Richards A, Quasnichka S, Taylor JG (2006) Population size and structure of whale sharks *Rhincodon typus* at Ningaloo Reef, Western Australia. Mar Ecol Prog Ser 319:275–285

Mizroch SA, Herman LM, Straley JM, Glockner-Ferrari DA, Jurasz C, Darling J, Cerchio S, Gabriele CM, Salden DR, Von Ziegesar O (2004) Estimating the adult survival rate of central north Pacific humpback whales (*Megaptera novaeangliae*). J Mammal 85(5):963–972

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Oosterhout CV, Hutchinson W, Wills D, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear population. Mol Ecol 3:489–495

Paetkau D, Waits LP, Clarkson PL, Craighead L, Vyse E, Ward R, Strobeck C (1998) Variation in genetic diversity across the range of North American brown bears. Conserv Biol 12:418–429

Pardini AT, Jones CS, Scholl MC, Noble LR (2000) Isolation and characterization of dinucleotide microsatellite loci in the Great White Shark, *Carcharodon carcharias*. Mol Ecol 9(8):1176–1178

Pardini AT, Jones CS, Noble LR, Kreiser B, Malcolm H, Bruce BD, Stevens JD, Cliff G, Scholl MC, Francis M, Duffy CAJ, Martin AP (2001) Sex-biased dispersal of great white sharks. Nature 412:139–140

Rousset F (2007) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Resources 8(1):103–106

Rowat D, Meekan MG, Engelhardt U, Pardigon B, Vely M (2007) Aggregations of juvenile whale sharks (*Rhincodon typus*) in the Gulf of Tadjoura, Djibouti. Environ Biol Fish 80:465–472. doi:10.1007/s10641-006-9148-7

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Press, Cold Spring Harbor

Schrey W, Heist EJ (2002) Microsatellite markers for the shortfin mako and cross-species amplification in lamniformes. Cons Genet 3:459–461

Sims DW, Speedie CD, Fox AM (2000) Movements and growth of a female basking shark re-sighted after a three year period. J Mar Biol Ass UK 80:1141–1142

Stevick PT, Palsbøll PJ, Smith TD, Bravington MV, Hammond PS (2001) Errors in identification using natural markings: rates, sources, and effects on capture–recapture estimates of abundance. Can J Fish Aquat Sci 58:1861–1870

Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. Mol Ecol Notes 2:377–379. doi:10.1046/j.1471-8286.2002.00228.x

Van Tienhoven AM, Den Hartog JE, Reijns RA, Peddemors VM (2007) A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*. J Appl Ecol 44:273–280. doi:10.1111/j.1365-2664.2006.01273.x

Vincent C, Fedak MA, McConnell BJ, Meynier L, Saist-Jean C, Ridoux V (2005) Status and conservation of grey seal, *Halichoerus grypus*, in France. Biol Conserv 126:62–73. doi:10.1016/j.biocon.2005.04.022

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity amond genotypes in natural populations: cautions and guidelines. Mol Ecol 10:249–256

Wilson B, Hammond PS, Thompson PM (1999) Estimating size and assessing trends in a coastal bottlenose dolphin population. Ecol Appl 9(1):288–300