<table>
<thead>
<tr>
<th>Title</th>
<th>Phylogeography, population structure, abundance and habitat use of bottlenose dolphins, Tursiops truncatus, on the west coast of Ireland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Nykanen, Milaja</td>
</tr>
<tr>
<td>Publication date</td>
<td>2016</td>
</tr>
<tr>
<td>Type of publication</td>
<td>Doctoral thesis</td>
</tr>
<tr>
<td>Rights</td>
<td>© 2016, Milaja Nykanen. <a href="http://creativecommons.org/licenses/by-nc-nd/3.0/">http://creativecommons.org/licenses/by-nc-nd/3.0/</a></td>
</tr>
<tr>
<td>Embargo information</td>
<td>Restricted to everyone for one year</td>
</tr>
<tr>
<td>Embargo lift date</td>
<td>2018-03-24T11:52:57Z</td>
</tr>
<tr>
<td>Item downloaded from</td>
<td><a href="http://hdl.handle.net/10468/3828">http://hdl.handle.net/10468/3828</a></td>
</tr>
</tbody>
</table>

Downloaded on 2018-12-18T04:32:33Z
# Table of Contents

Acknowledgements ................................................................................................................. 1

Declaration .................................................................................................................................. 3

**Chapter 1: General Introduction** .......................................................................................... 7

1.1 Taxonomy and phylogenetics of bottlenose dolphins ......................................................... 8

1.2 Biogeography .................................................................................................................... 13

1.3 Distribution and habitat use ........................................................................................... 14

1.4 Social structure ............................................................................................................... 16

1.5 Population structure of bottlenose dolphins in the North Atlantic Ocean .............. 17

1.6 The use of photo-identification and mark-recapture methods in the study of bottlenose dolphins .............................................................................................................. 20

1.7 Conservation and management of bottlenose dolphins ................................................. 22

1.8 Background to the project and main aims of the study ............................................. 24

References .................................................................................................................................. 29

**Chapter 2: Modelling the post-glacial colonisation of the northern extreme of the species range in the bottlenose dolphin *Tursiops truncatus* (Montagu, 1821) ................................................................. 42

2.1. Introduction ..................................................................................................................... 44

2.2. Materials and Methods ................................................................................................. 46

DNA extraction, amplification and sequencing .................................................................. 48

Phylogenetic and coalescent analyses ............................................................................... 49

Phylogeographic methods ................................................................................................. 52

Models of suitable habitat ................................................................................................. 53

2.3. Results ............................................................................................................................. 56

Calibration of root node age.............................................................................................. 56

Time-calibrated phylogeny of North Atlantic bottlenose dolphins .................................. 58

Models for suitable habitat and phylogeography ........................................................... 61
Chapter 2: Appendices .............................................................................................................. 82

Chapter 3: Quantifying dispersal between marine protected areas by a highly mobile species, the bottlenose dolphin, *Tursiops truncatus* ............. 87

3.1 Introduction ......................................................................................................................... 89

3.2 Materials and Methods ................................................................................................. 92

Photo-identification surveys and photograph selection ............................................... 92
Skin tissue sample collection and analysis............................................................................ 92
DNA extraction, PCR amplification and genotyping ......................................................... 95
Individual assignment tests ................................................................................................. 96
Migration rates ...................................................................................................................... 99
Effective population size ..................................................................................................... 99
Analyses of social structure and site fidelity ................................................................. 100
Relatedness, associations and spatial overlap ............................................................ 101

3.3 Results ............................................................................................................................... 102

Individual assignment tests ............................................................................................. 102
Population differentiation and effective population size ............................................. 109
Sex-biased dispersal and migration rates ...................................................................... 111
Social structure and site fidelity .................................................................................... 113
Relatedness, spatial overlap and associations ............................................................ 117

3.4 Discussion ......................................................................................................................... 118

References .............................................................................................................................. 123

Chapter 3: Appendices .......................................................................................................... 132

Chapter 4: Using Bayesian inference with a multi-site mark-recapture model to estimate the abundance of a mobile population of bottlenose dolphins, *Tursiops truncatus*, on the west coast of Ireland .................. 139

4.1. Introduction ....................................................................................................................... 141
Chapter 6: Concluding discussion ................................................................. 197

Commonly used methods in conservation and management in the context of the present study ................................................................. 197

Areas for future research ........................................................................... 205

Utility of Marine Protected Areas in the conservation of marine mammals .... 207

References .................................................................................................... 210

Chapter 6: Appendices ................................................................................ 217
Acknowledgements

I wish to thank my supervisors for all your help and support over the course of this project; Emer, thank you so much for caring, I feel like I have learned a lot from you and I could always turn to you for support if there were any problems. Andy, you have been an inspiring supervisor and a great friend. Thanks for all the support and the push you have given me, I could not have done this without you. Thank you Simon especially for great field work and biopsy training, it has been a pleasure to have been able to learn from you.

This PhD was funded by the Crawford-Hayes studentship. The funding for the field work was provided by the National Parks and Wildlife Service, under the Department of Arts, Heritage and Gaeltacht, Ireland. The biopsy samples collected during this study were obtained under permits from the Irish National Parks and Wildlife Service, Department of Arts, Heritage and the Gaeltacht, and the Irish Medicines Board, Health Products Regulatory Authority. Some of the samples from stranded animals were obtained from the Irish Cetacean Genetic Tissue Bank, Natural History Museum, Dublin. Big thank you also for everyone who provided tissue samples; Marie Louis, Willy Dabin, Vincent Ridoux, Andrew Brownlow, Nick Davison and Rob Deaville. Thank you to Vincent Biard, Pavel Gol’din and Morten Tange Olsen for providing the ancient sequences, to Kristin Kaschner and Christina Garilao for providing the AquaMaps output, and to Nathan Wales for the help with the sequencing. Thank you also to Henry Van der Es and Klaas Post Natural History Museum in Rotterdam for letting me sample the museum collection of subfossils. Big thanks to Enrico Pirotta for the help with modelling the C-POD data.

I would also like to thank Oliver Ó Cadhla and Ferdia Marnell for their help and support during the project. Big thanks to Eileen Dillane for the lab work and to everyone who helped with biopsy sample and data collection, including Rúadhán Ó Ceallaigh, Marie Kearns, Ross Culloch, Allen Whitaker, Barry McGovern, Lochie O’Kelly, Brian and Cyndi from St John’s Point, Machiel Oudejans, Damian Haberlin, Martha Gosch, Róisín Pinfield, Claudia Melville, the Irish Whale and Dolphin Group (especially Caroline Tuffy), Killary Cruises, skippers of the Pirate Queen, Jarlath Hession, Shane Bisgood, John Britain, Máirtín Ó Meallaigh, Selkie...
sailing, Killybegs Coast Guard, Paddy Byrne, Robert Beirne, and all the lovely people of Connemara, Mayo and Donegal.

I would also like to thank all my colleagues and friends in the School of BEES, UCC, and MaREI who have made the last few years in Cork so enjoyable. A special thanks to Bob Kelly, Martha Gosch, Ash Bennison and Gavin Arneill for putting up with me during stressful times, you guys rock! Finally, I would like to thank my family and friends in Finland for their continued encouragement and support.
Declaration

I declare that this thesis has not been previously submitted as an exercise for a degree at National University of Ireland, Cork, or any other university and I further declare that the work embodied in it is my own, or else noted.

Milaja Nykänen
Glossary of genetic terms used in the thesis

**Allelic dropout**: a source of missing data in microsatellite genotypes, in which one or both allelic copies at a locus fail to be amplified by the polymerase chain reaction.

**Biogeography**: the study of the distribution of species in geographic space and through (geological) time.

**Cladogenesis**: the formation of a new group of organisms or a higher taxon by evolutionary divergence from a “parent” taxon.

**Coalescent theory**: a stochastic model of population genetics that relates genetic diversity in a sample to demographic history of the population from which it was taken.

**Complete lineage sorting**: a perfect segregation of all alleles into all lineages.

**DNA barcoding**: a taxonomic method that uses a short genetic marker in an organism's DNA to identify it as belonging to a particular species.

**DNA library**: a collection of DNA fragments.

**DNA sequencing**: the process of determining the order of nucleotides within a DNA molecule.

**Ecotype**: Here, defined as groups of populations that differ across geographic space in genetic (e.g. allele frequencies) and other (e.g. ecological, morphological, physiological) traits.

**Effective population size** (*N_e*): the number of individuals that an idealised population (with random mating, simultaneous birth of each generation, constant population size, equal number of offspring per parent) would need to have, in order for some specified quantity of interest to be the same in the idealised population as in the real population.

**Fixation index** (*F_{ST} and F_{IS}*)**: a measure of population or individual differentiation due to genetic structure (*F_{ST}* ) or inbreeding (*F_{IS})*.
**Founder effect**: a non-random sampling that can exclude alleles from a new subpopulation by chance.

**Founder effect**: the loss of genetic variation that occurs when a new population is established by a very small number of individuals from a larger population (a founder event).

**Genetic drift**: random sampling of allele frequencies in a population.

**Haplotype**: a group of genes in an organism that are inherited together from a single parent. A haplogroup is a group of similar haplotypes that share a common ancestor with a single nucleotide polymorphism mutation.

**Hardy-Weinberg equilibrium (HWE)**: Principle stating that the genetic variation in a population will remain constant from one generation to the next in the absence of disturbing factors.

**Incomplete lineage sorting**: Under incomplete lineage sorting in population genetics, the coalescence time of genes to the same common ancestor and speciation time are different. Also referred to as deep coalescence when gene coalescence times are much older than species divergence times.

**Introgression**: Gene flow from one species into the gene pool of another by the repeated backcrossing of an interspecific hybrid with one of its parent species.

**Linkage disequilibrium**: the non-random association of alleles at different loci.

**Microsatellite DNA**: a sequence of repetitive DNA in which certain motifs (ranging in length from 2–5 base pairs) are repeated.

**Models of DNA sequence evolution**: To account for gene effects in the data set, each gene can be assigned a parameter that describes its substitution rate, e.g. GTR, HKY.

**Multiplex polymerase chain reaction (Multiplex PCR)**: performing many separate PCR reactions all together in one reaction.

**Null allele**: a mutant copy of a gene at a locus that completely lacks that gene's normal function.

**Operational Taxonomic Unit (OTU)**: DNA sequences can be clustered based on their similarity to one another according to similarity threshold set by the researcher.
**PCR**: Polymerase chain reaction (PCR) is a technique used in molecular biology to amplify a single copy or a few copies of a piece of DNA.

**Phylogenetic tree**, or **phylogeny**: a branching diagram displaying the inferred evolutionary relationships among taxons based upon similarities and differences in their physical or genetic characteristics.

**Phylogeography**: the joint phylogenetic relationships and geographic distributions of genetic lineages.

**Population bottleneck**: an abrupt reduction in population size due to environmental events or human activities.

**Purifying selection**: Many of the sequence polymorphisms that are seen among individuals of a population and in intraspecific comparisons are removed over evolutionary time due to the action of purifying selection or by random genetic drift.

**Relaxed molecular clock**: takes into account variability in substitution rates between lineages of the phylogeny.

**Strict molecular clock**: Nucleotide substitution rate model where the expected number of substitutions per year is constant regardless of which species' evolution is being examined.

**Tajima’s D**: a statistic that compares the average number of pairwise differences with the number of segregating sites.

**Time-dependency (in substitution rate estimation)**: Theory according to which nucleotide substitution rate estimates based on recent calibration points are much higher than those calibrated by older nodes.

**Tree-branching models**: the rate at which new branches are formed in a phylogenetic tree, e.g. Yule, birth-death model, coalescent model, etc.
Chapter 1: General Introduction

Bottlenose dolphins are among the most well-studied cetaceans in the world, with nearly 29,000 hits returned on Google Scholar using the search term “bottlenose dolphin”. The early research on bottlenose dolphins concentrated largely on the natural history and distribution of the species (e.g. Miller 1923), and these were followed by studies on the morphology, physiology and behaviour, where data was generally collected by killing free ranging animals or capturing live animals for display and research purposes (e.g. True 1890; Gunter 1942; Fetcher & Fetcher 1942; McBride & Hebb 1948; Lilly & Miller 1961; Lilly 1962, 1963). What makes bottlenose dolphins such a well-studied species? One possible reason may lie within their acclaimed intelligence and their abilities to exhibit complex social structures and behaviours otherwise found in the animal kingdom only within other delphinids, primates and elephants (e.g. Reiss & Marino 2000; Janik 2000; Connor 2000; Connor 2007; Möller et al. 2012). Other reasons for the appeal of this species among scientists may be their world-wide distribution (Fig. 1.1) and the fact that the proximity of some coastal bottlenose dolphin populations to human settlements makes them more available as a study species compared to some more elusive or hard to reach cetaceans. At the same time this proximity to land also makes them more vulnerable to human impacts – a key reason why coastal populations have become the focus of conservation efforts, requiring management and monitoring strategies to be put to place.

The fact that they spend their lives under water makes studying wild bottlenose dolphin populations difficult and presents several challenges as the data collection is usually constrained to good weather and light conditions. Nevertheless, the fact that individual dolphins can often be distinguished from unique sets of markings accumulated on their dorsal fins or bodies (e.g. Würsig & Würsig 1977; Würsig & Jefferson 1990), makes it possible to gather information on their abundance, movements and social structure through mark-recapture techniques. The methods relating to these topics are discussed more thoroughly in Chapters 3 and 4. However, whilst the resident coastal populations can be accessed relatively easily by using small vessels and sometimes even from land, collecting data on coastal individuals or populations that exhibit more unpredictable ranging patterns or whose distribution is concentrated to areas further offshore, presents even further challenges. In situations like these, alternative strategies, such as
the genetic sampling of stranded or bycaught individuals, can provide a way for gathering information on these populations (Chapters 2 and 3). In mark-recapture abundance estimation, on the other hand, alternative modelling approaches that allow for more flexible survey schemes can provide efficient and cost-effective alternatives to traditional models (Chapter 4). Furthermore, using a combination of different methodologies, such as a passive acoustic monitoring (PAM, Chapter 5) coupled with visual monitoring methods, can provide a way of monitoring habitat use of bottlenose dolphins and may even be used to direct (visual) survey effort into an appropriate area and season.

Figure 1.1 Map showing the world-wide distribution of *Tursiops truncatus*. With the exception of polar regions, the species is found throughout the world’s oceans. (IUCN 2012).

1.1 Taxonomy and phylogenetics of bottlenose dolphins

In short, the purpose of taxonomy is to provide a way to organise and summarise information on the relationships of organisms and to divide them into hierarchical
classes based on similarities and dissimilarities. This kind of classification can also reveal information on the evolutionary pathways along which present-day organisms may have developed from ancestral forms. Before the introduction of molecular genetics and DNA sequencing technologies, studies resolving the taxonomy of organisms within the family Delphinidae were primarily done by careful examination and measurement of morphological features, with particular focus on the morphology of the skull (reviewed by LeDuc et al. 1999). Even today morphometric methods are used to support the phylogenetic relationships within the delphinid family (e.g. Amaral et al. 2009). Recently, methods that combine information on fossil morphometrics and genetic data into ‘total-evidence dating’ phylogenetic models have been developed (Ronquist et al. 2012).

Due to its clock-like nucleotide substitution rate and the fact that it is expected to conserve and accumulate mutations over time (Ho et al. 2008), the reconstruction of the phylogenetic history of organisms has primarily been done by sequencing parts of the maternally inherited mitochondrial genome. The first studies resolving the phylogenetic history of delphinids used only short sequences of the genome, such as the control region and cytochrome b gene in the genus Delphinus (Rosel et al. 1994) and family Phocoenidae (Rosel et al. 1995), and at a wider taxonomic level of the subfamily Delphininae (LeDuc et al. 1999), but during recent years, the emergence of Next Generation Sequencing (NGS) technologies has vastly increased the amount of genetic data that can be generated for phylogenetic studies. This combined with the development of phylogenetic statistical tools (see review by Yang & Rannala 2013), has improved the ability to find phylogenetic differences and increased the resolution in species divergence estimation. For example, Steeman et al. (2009) used six mitochondrial and nine nuclear genes combined with fossil calibrations to reconstruct the phylogenetic history of cetaceans (Fig. 1.2), and within the same year, a study by McGowen et al. (2009) used an impressive set of mitochondrial genomes and nuclear genes to build a comprehensive, time-calibrated phylogenetic tree of the same taxonomic group.

Previous phylogenetic studies, including the two mentioned above, have all suggested a relatively recent (beginning ~10 million years ago) and rapid radiation within Delphinidae (Steeman et al. 2009; McGowen et al. 2009; Kingston et al. 2009; Xiong et al. 2009), hypothesized to be linked to periods of environmental fluctuations
(McGowen et al. 2009, Steeman et al. 2009: Vilstrup et al. 2011), which has led to uncertainties in the phylogenetic relationships within the family. Vilstrup et al. (2011) aimed to resolve these uncertainties by increasing the amount of mitogenome sequences from this family, including several sequences from within the same species, followed by Amaral et al. (2012) who constructed a species tree from both mitochondrial and nuclear data using coalescent analysis (Fig. 1.3).
Figure 1.2 Phylogenetic tree of the order Cetacea taken from Steeman et al. (2009). Note the rapid radiation observed in the family Delphinidae.
Figure 1.3 Phylogenetic trees of the superfamily Delphinoidea taken from Vilstrup et al. (2011) (on the left) and from Amaral et al. (2012) (on the right).
The taxonomy within the genus *Tursiops* is similarly unclear, and depending on the DNA markers and phylogenetic method used, contradicting topologies have been produced (LeDuc *et al.* 1999, Wang *et al.* 1999; Vilstrup *et al.* 2011; Amaral *et al.* 2012). In addition to the common bottlenose dolphin, *Tursiops truncatus* (Montagu 1821), Indo-Pacific bottlenose dolphin, *Tursiops aduncus*, has been recognized as a separate species by a number of international organisations, including the International Union for Conservation of Nature (IUCN) and the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) based on phylogenetic and morphological evidence (Curry & Smith 1997; LeDuc *et al.* 1999, Wang *et al.* 1999). This was followed by another species designation of *T. australis*, the Burrunan dolphin, based on differences in morphology, colouration, cranial characters and genetic evidence (Charlton-Robb *et al.* 2011). However, the validity of the latter species is still subject to debate and not yet fully recognised (Committee on Taxonomy 2014).

Following a debate within CITES whether a putative Black Sea subspecies of bottlenose dolphin, *Tursiops truncatus ponticus*, should be given a full species status, a specialised Workshop on Cetacean Systematics was held to review cetacean taxonomy (Reeves *et al.* 2004). As a result of this workshop, it was decided that for species delineation, both morphological and genetic data would have to be consistent to prove “irreversible reproductive isolation” (Reeves *et al.* 2004). A key question is thus whether the genus *Tursiops* should undergo further division with the separation of the two forms (or ‘ecotypes’) in the Western North Atlantic, an offshore and a coastal form based on morphology and ecological and genetic markers (Duffield *et al.* 1983; Hersh & Duffield 1990; Mead & Potter 1995; Leduc & Curry 1997, Hoelzel *et al.* 1998).

1.2 Biogeography

Moura *et al.* (2013) analysed an impressive set of complete mitochondrial genomes, which included samples collected from Australia, Africa, North Pacific, North Atlantic, Mediterranean and Black Sea. The results suggested that the extant members of the genus *Tursiops* originated from Australasia from lineages occupying coastal habitats. The authors also hypothesised that the expansion of the lineages to the
Atlantic occurred via Indo-Pacific coastal habitats followed by the colonization of the pelagic area and the delineation of the *T. truncatus* species. *T. truncatus* then regressed back to the ancestral coastal state and are represented, among others, by the coastal populations on the West Coast of North America (Moura *et al.* 2013).

1.3 Distribution and habitat use

Resources in natural ecosystems are variable and usually patchily distributed in both time and space. It is thus not surprising that most marine predators are not randomly or uniformly distributed, but that their distribution can be driven by a combination of abiotic or biotic factors, such as depth, temperature, ocean currents and fronts, benthic type, bathymetry, prey distribution, presence of predators or conspecifics and anthropogenic factors (e.g. Ford *et al.* 1998; Saulitis *et al.* 2000; Heithaus & Dill 2002; Ingram & Rogan 2002; Johnston *et al.* 2002; Olesiuk *et al.* 2002; Williams *et al.* 2002; Embling *et al.* 2010; Sveegard *et al.* 2012; Booth *et al.* 2013; Pirotta *et al.* 2011; Scales *et al.* 2014). Bottlenose dolphins are top marine predators with a world-wide distribution extending from tropical to temperate waters (Leatherwood and Reeves 1990; Connor *et al.* 2000, see Fig. 1.1). Fine-scale habitat use of bottlenose dolphins has been linked to a variety factors. For example, Wilson *et al.* (1997) found that dolphins in the Moray Firth, Scotland, occurred mostly in areas with deep channels associated with strong tidal currents. Bottlenose dolphins were present in the bay year-round but the use of different parts of the bay changed seasonally with the number of individuals increasing in the summer months (Wilson *et al.* 1997) and during flood tides (Mendes *et al.* 2002). Further, video surveillance in these channels revealed that the dolphins were seen most frequently within the deepest waters and areas of steep seabed gradients (Hastie *et al.* 2003). Subsequently, Hastie *et al.* (2004) found evidence that the dolphins were engaged in feeding behaviour in these high use areas, with a correlation between the occurrence of feeding and increasing bathymetry (Hastie *et al.* 2004). Within the Shannon estuary, Ireland, bottlenose dolphins were also encountered in deeper areas with steep sloping benthic topography rather than in shallower areas with lower benthic gradients, even though individual differences existed in habitat use (Ingram & Rogan 2002).
In environments with less local variation caused by benthic topography, such as the open ocean pelagic zone, seasonal effects can have a stronger influence on habitat use of marine mammals. For example, Bearzi et al. (2008) found several environmental covariates explaining the presence of bottlenose dolphins but the effect varied depending on the season.

In addition to the distribution of resources and conspecifics, the presence of predators can have an effect on the habitat use of animals. Indo-Pacific bottlenose dolphins (*T. aduncus*) in Shark Bay, western Australia were found to utilize the food-rich shallow habitats more during winter months when predation risk from tiger sharks (*Galeocerdo cuvier*) was lower, moving towards less productive deeper areas for summer months when the predation risk was increased (Heithaus & Dill 2002). In addition, the distribution of dolphins was found to be approximately proportional to prey density when shark abundance was low but at high shark abundance the dolphin distribution changed towards safer, deeper, but less productive waters (Heithaus & Dill 2002). The presence of another shark predator, bull shark (*Carcharhinus leucas*), during the summer months in the inshore waters of the Gulf of Mexico may have driven the distribution of bottlenose dolphins into protected bays (Wells & Scott 2002).

Anthropogenic disturbance can have an effect on the habitat use of marine mammals and it can even cause permanent displacement from an area. In New Zealand, for example, bottlenose dolphins have been encountered less in Milford Sound, as a result of increased boat traffic (Lusseau 2005). Similarly, the abundance of Indo-Pacific bottlenose dolphins in Shark Bay has decreased in areas with dolphin-watching boats (Bejder et al. 2006). Kuningas et al. (2013) found the presence of killer whales, *Orcinus orca*, to be negatively affected by naval sonar activity but this effect was masked by herring (*Clupea harengus*) abundance, the main factor affecting killer whale presence in that area. Further, Miller et al. (2015) found naval sonar signals to affect the diving and echolocation behaviour of northern bottlenose whales *Hyperoodon ampullatus*; the animals ceased their echolocation during the exercise and one animal that was tracked avoided the sound source by changing direction and swimming away from it.
1.4 Social structure

Bottlenose dolphins generally live in fluid “fission-fusion” societies (Connor et al. 2000), which means that dolphins usually form small social groups whose composition can change rapidly within the scale of a few hours (see page 109 in Connor et al. 2000). The fluidity of dolphin social networks has recently been linked to the low cost of locomotion (Randić et al. 2012). Gender can also influence the associations of dolphins, however, this does not apply to every population. In Shark Bay, Australia, Smolker et al. (1992) observed long term associations in Tursiops spp, generally between members of the same sex with the strongest associations between a mother and her offspring or between two males. These first-order male subgroups in turn had moderate associations with other male first-order subgroups thus forming second-order alliances. Similarly, females preferred to associate with certain other females forming a network where all females were interconnected through a chain of consistent associates, but in general, the associations were less stable (Smolker et al. 1992). Connor et al. (2000) hypothesized that the male-male alliances in Shark Bay are formed due to the need for co-operation in herding females. The social structure and stability of these first-, second- and even third-order alliances in Shark Bay have recently been re-examined and reviewed (Connor et al. 2011; Connor & Krützen 2015), and the results have suggested that second order alliances may persist for up to 20 years, highlighting the importance of long term studies in understanding dolphin societies. These stable alliances may be formed in order to increase individuals’ inclusive fitness by kin selection; for example, Krützen et al. (2003) found that the males in stable first- and second-order alliances were often strongly related. Similarly, Frere et al. (2010a) found that female bottlenose dolphins preferred to associate with closely related individuals than expected by chance. Close social bonds have been found to correlate with increased calving success thus promoting individual’s fitness (Frere et al. 2010b).

Strong male-male associations between bottlenose dolphins were also observed in Sarasota, Florida, but co-operative second-order alliances were not recorded (Connor et al. 2000). In contrast, no strong sex-specific alliances were found in the Moray Firth (Wilson 1995) or in the Shannon estuary in Ireland (Foley et al. 2010), and this has been suggested to be caused by local site-specific differences in resource competition and/or the reduced risk of predation (Wilson 1995). Sex-specific alliances were not
observed amongst the dolphins in Doubtful Sound, New Zealand, but some male–female associations were stable over the course of several years (Lusseau et al. 2003). Specialised foraging techniques, which may promote the formation of social groups and thus have important ecological and evolutionary consequences contributing to niche specialisation and possibly even to genetic structuring (Krützen et al. 2014; Kopps et al. 2014), have also been observed in bottlenose dolphin communities (Mann & Sargeant 2003). For example, some social groups in Shark Bay are known for their tool use, i.e., “sponging” (Smolker et al. 1997; Mann et al. 2008; Krützen et al. 2005, 2014) and to a lesser extent, “conching” (Allen et al. 2011). Strand feeding (e.g. Duffy-Echevarria et al. 2008; Jimenez et al. 2015) and beach hunting (Sargeant et al. 2005) have been observed in dolphin communities in different parts of the world along with feeding associated with trawlers (e.g. Broadhurst 1998; Chilvers & Corkeron 2001; Gonzalvo et al. 2008). Interestingly, Chilvers and Corkeron (2001) found dolphins in Moreton Bay, Australia, feeding from trawling nets to almost exclusively associate with other “trawler feeders” thus forming a separate community from other individuals who did not use this foraging tactic. However, this social separation disappeared when trawling was banned in the bay, being replaced by a more fluid social structure with more associations occurring between all the dolphins (Ansmann et al. 2012), which indicates that social clusters can be fluid and may form as a response to a newly available resource and dissipate when this resource is removed. Another example of social groups forming based on specialized foraging techniques include the bottlenose dolphins in Laguna, Brazil, where certain bottlenose dolphins have learned to cooperate with fishermen, seemingly driving mullet (Mugil spp.) into their nets and catching the escaping fish (Simões-Lopes et al. 1998; Daura-Jorge et al. 2012).

1.5 Population structure of bottlenose dolphins in the North Atlantic Ocean

The lack of physical barriers in the marine environment combined with the fact that many marine organisms have good capabilities for dispersal, should, in theory, mean high gene flow and lack of speciation in marine organisms. This is the case with most marine species, as very little genetic differentiation is usually observed even with distances of thousands of kilometres (reviewed by Palumbi 1994). In highly mobile organisms, such as birds and mammals, gene flow is often reduced due to socio-
ecological, behavioural and/or environmental factors that can lead to divergence of populations. For example, female philopatry and sex-biased dispersal, combined with short dispersal distances shown by the less philopatric males may have led to fine-scale population structuring of co-operatively breeding birds (Temple et al. 2006). Similarly, long-term fidelity to natal social clusters shown by both sexes, possibly enforced by differences in foraging strategies required to exploit resources in varying environments are thought to have led to fine-scale population structuring of two resident populations of false killer whales (*Pseudorca crassidens*) around Hawaii (Martien et al. 2014). Wolf et al. (2008) also found natal philopatry among breeding female Galápagos sea lions (*Zalophus wollebaeki*) and hypothesized habitat specialization as one of the driving forces behind genetic divergence. High levels of relatedness between close associates may have further driven fine-scale population structuring of some species (e.g. Temple et al. 2006; Iacchei et al. 2013), even when accounting for spatial proximity (Podgórska et al. 2014), thus reinforcing interactions among related individuals and encouraging philopatry especially when food sources are limited or patchy. Similarly, many marine organisms show population structure over small geographic scales (see review by Palumbi 1994; Bierne et al. 2003). This can lead to severe loss of heterozygosity over time due to genetic drift and inbreeding, especially if populations are small and isolated (Lacy 1987). Possible drivers of this fine-scale population structure include existence of ocean fronts (White et al. 2010), historic and oceanographic influences (Woodall et al. 2015), isolation by distance, historical founding events, complex social interactions, natal philopatry, and development of foraging specializations and habitat preference possibly leading to adaptive isolation (e.g. Natoli et al. 2005; Krützen et al. 2004; Rosel et al. 2009). In a recent review, Möller (2012) suggests that environment type often dictates the social bonds observed in populations, with more female-biased philopatry found among small delphinids inhabiting shallow inshore environments with more predictable resources when compared to those inhabiting offshore deeper environments where food availability is less predictable.

The bottlenose dolphin shows hierarchical population structure throughout its worldwide range, with the greatest divergence found between pelagic and coastal populations (Curry & Smith 1997; Hoelzel et al. 1998; Louis et al. 2014a; Lowther-Thieleking et al. 2015). This is often accompanied by ecological and/or morphological
differences (Duffield et al. 1983; Hersh & Duffield 1990; Hoelzel et al. 1998; Louis et al. 2014a; Natoli et al. 2004) and a division to offshore and inshore forms, or *ecotypes*, has been suggested based on these differences in the Northwest Atlantic (Duffield et al. 1983; Mead & Potter 1995) and more recently, also in the Northeast Atlantic (Louis et al. 2014a). Moreover, Hoelzel et al. (1998) analysed nuclear microsatellites from nearshore and offshore bottlenose dolphins from the western North Atlantic (WNA), the Bahamas and Africa and found that the level of genetic variation among the nearshore dolphins was reduced compared with the offshore population in the WNA. Greater genetic diversity among the North Atlantic offshore bottlenose dolphins has since been found in other studies (e.g. Natoli et al. 2004; Tezanos-Pinto et al. 2009; Mirimin et al. 2011; Louis et al. 2014a, see also Chapters 2 and 3). However, the division to inshore and offshore forms may not necessarily apply in other parts of the world; Hoelzel et al. (1998) found all the nearshore African haplotypes to cluster together with the WNA offshore population.

Parsons et al. (2002) compared microsatellite markers sampled from bottlenose dolphins around the UK and found significant levels of genetic differentiation between the northeast of Scotland (the Moray Firth) population and the west coast of Scotland populations. In fact, the NE Scotland dolphins appeared to be more closely related to the bottlenose dolphins occupying Cardigan Bay, Wales, than to the ones found in the Sound of Barra on the Scottish west coast, despite the larger geographic distance (Parsons et al. 2002). Mirimin et al. (2011) analysed biopsy samples from bottlenose dolphins along the west coast of Ireland and found that the samples collected in Connemara (Co. Galway) and Mayo (Co. Mayo) belong to the same genetic population. These samples differed significantly from samples collected in the Shannon estuary (~150km south of Connemara) with both populations showing low genetic diversity (Mirimin et al. 2011). However, most samples collected from stranded dolphins along the Irish west coast showed much greater genetic diversity compared to the coastal populations, and Mirimin et al. (2011) hypothesised that these samples likely originated from a third putative offshore population inhabiting the waters between the coast and continental shelf. The dolphins using the coastal waters of western Ireland appear to be also socially distinct from dolphins using offshore waters (Oudejans et al. 2015).
Nichols et al. (2007) suggested that coastal bottlenose dolphins in the NE Atlantic may be part of a wider meta-population based on reduced, but ongoing gene flow among small local populations that seem to be dependent on local habitat patches (Natoli et al. 2005; Sellas et al. 2005), and this was further supported by Louis et al. (2014a) whose study clustered all the coastal samples collected in Ireland and the UK into a single ‘Coastal North’ population. In a further study, also based on neutral genetic markers, Louis et al. (2014b) suggested that the broad-scale division to ‘coastal’ and ‘pelagic’ ecotypes in the NE Atlantic may reflect a historical divergence of populations followed by colonisation events of the coastal populations into available inshore habitats from an oceanic source population after the Last Glacial Maximum, LGM (Louis et al. 2014b). Divergence between the ‘Coastal North’ and the ‘Pelagic Atlantic’ populations was estimated to date to ca. 10,320 years before present (yBP) and the colonisation was suggested to have occurred via a small number of founding individuals (Louis et al. 2014b). The end of the last glacial period in the Northern Hemisphere is thought to have had a major impact on genetic diversity of organisms (Hewitt 1999, 2000), and post-glacial colonisations such as the above would have been a common pattern during post-glacial periods, leaving a genetic signature in present day populations (Hewitt 2000). Furthermore, Moura et al. (2013) found the differentiation of the coastal and pelagic WNA ecotypes coinciding with periods of fast climatic change during the Eemian and Holocene periods, and suggested that climatic oscillations may be involved in the diversification of bottlenose dolphins in this area (Moura et al. 2013). Conversely, Louis et al. (2014b) concluded that the divergence of the ‘Coastal North’ and ‘Coastal South’ populations (ca. 2,560 yBP) was not linked to any particular climatic event.

1.6 The use of photo-identification and mark-recapture methods in the study of bottlenose dolphins

Understanding the distribution, movements and abundance of cetacean populations currently relies largely on our ability to recognise individuals from unique markings. Photo-identification is a technique used to identify individual animals from photographs using naturally occurring markings as distinctive characters (see Fig. 4.2 in Chapter 4). This method was first used to identify individual bottlenose dolphins by Caldwell (1955), Irvine & Wells (1972) and Würsig & Würsig (1977), since most
bottlenose dolphins accumulate distinctive notches and scars on their dorsal fins and bodies through social interactions with conspecifics.

Mark-recapture, or capture-recapture methods in abundance estimation of closed and open populations are mostly different variations of the Petersen estimator (Seber 1982, 1992). With ‘population closure’ it is assumed that the population is stable, with net migration and the difference between births and deaths both equalling zero; needless to say that this assumption is rarely fulfilled in natural populations. The Petersen estimator, \( \hat{N} \), is obtained by equating the proportion of initially marked and released individuals \( (m_1) \) to the total number of animals in the population \( (N) \) with the proportion of marked animals \( (m_2) \) to the number of unmarked and marked animals in the second sample \( (n_2) \)

\[
\hat{N} = \frac{m_1 n_2}{m_2}
\]

If more than two samples are collected, the unmarked animals can be marked in the second sample and subsequently all the animals released again. This sequential mark and release process will lead to individual capture histories (reviewed by Schwartz & Seber 1999).

Mark-recapture abundance models are inevitably associated with uncertainty resulting from heterogeneity in the capture, survival and sighting probabilities (Scwartz & Seber 1999), and several modifications to the models have been developed to quantify this heterogeneity (e.g. Seber 1992). Possibly the most frequently used method to estimate bottlenose dolphin abundance has been \( M_{\text{th}} \) by Chao et al. (1992) due to its good performance in abundance estimation of populations with reasonable amounts of heterogeneity. Log-linear models (Fienberg 1972) are also particularly useful for modelling both heterogeneity in capture probabilities and their dependencies between sampling occasions. In general, due to the wide variety of available mark-recapture models, the choice for the most suitable model should ideally depend on the characteristics of the data collected (Durban et al. 2005).

Bayesian statistics, as opposed to traditional frequentist Maximum Likelihood (ML) based estimation, have recently become more commonly used in mark-recapture abundance estimation (e.g. Mäntyniemi & Romakkaniemi 2002; Michielsens et al. 2006), and they have especially been applied to abundance estimation of a variety of
Bayesian methods are particularly useful when data are sparse (e.g. Durban et al. 2005; Royle et al. 2007; Cheney et al. 2013) due to the possibility of using Monte Carlo methods to sample from the posterior distribution (Scwartz & Seber 1999). This makes them highly applicable to cetacean abundance studies that are often burdened by small data sets due to challenges involved in data collection resulting from, for example, poor weather conditions or unpredictable occurrence and elusive behaviour of the study species.

1.7 Conservation and management of bottlenose dolphins

On a global scale, common bottlenose dolphins, are considered as of ‘least concern’ by the International Union for Conservation of Nature (IUCN) Red List of Threatened Species (Hammond et al. 2012). Subspecies T. truncatus ponticus in the Black Sea is listed as ‘endangered’ and the T. truncatus populations occurring in the Mediterranean Sea ‘vulnerable’ (Birkun 2012; Bearzi et al. 2012). Bottlenose dolphins are also listed in Annex II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), which prohibits their commercial trade in countries that have signed the treaty.

In European waters, bottlenose dolphins are protected through the Habitats Directive (92/43/EEC), and like all cetaceans, they are listed in Annex IV of the Habitats Directive necessitating ‘strict protection’ for such species. In addition, bottlenose dolphins are listed in Annex II of the Habitats Directive; this means that the Member States are required to designate Special Areas of Conservation (SACs) as part of a European strategy to maintain or restore a favourable conservation status for the species (Natura 2000). Further, as top predators, bottlenose dolphins are considered as one of the indicator species for ‘good environmental status’ in coastal waters in the Marine Strategy Framework Directive (MSFD, 2008/56/EC). The aim of the MSFD is to protect the European marine environment by applying a wholesome ecosystem-based approach to the management of human activities in these waters whilst encouraging sustainable use of the environment.

Despite the lack of consensus among researchers on what constitutes a population (Waples & Gaggiotti 2006), defining populations either as Management Units (MUs)
or Evolutionarily Significant Units (ESUs) (Moritz 1994), is relevant when setting strategies for conservation and management. Moritz (1994) defined MUs as “population units with significant divergence of allele frequencies at nuclear or mitochondrial loci, regardless of the phylogenetic distinctiveness of the alleles”. ESUs, on the other hand, were defined by Moritz (1994, 2002) as units arising from “historical population structure rather than current adaptation that are reciprocally monophyletic for mitochondrial DNA and show significant divergence of allele frequencies at nuclear loci”. Crandall et al. (2000) added the importance of ecological exchangeability to the equation stating that “the rejection of, or failure to reject, (genetic or ecological) exchangeability forms the foundation of population distinctiveness and management practices”. However, Waples and Gaggiotti (2006) and Palsbøll et al. (2006) criticised the use of the statistical criterion in these definitions stating that population structure may go undetected when population genetic divergence is low and has occurred relatively recently, since the statistical power to detect population structure is a function of the amount of data and thus correlates to the number of loci used and samples analysed. In these cases, using ‘kinship-based’ approaches and/or other methods in quantifying recent (past few generations) migration among putative populations may offer a solution to the delineation of MUs (Waples & Gaggiotti 2006; Palsbøll et al. 2010).

Bottlenose dolphin populations using coastal environments are at particular risk of exposure to a number of anthropogenic threats which may directly impact individuals, for example through disturbance or damage to health and to the overall functioning of the coastal ecosystems upon which they depend. The sensitivity of bottlenose dolphins to these threats is exacerbated by their position as an apex predator and also by their low reproductive rates (Connor et al. 2000; Quick et al. 2014). The main threats in coastal environments include pollutants such as xenobiotic chemicals (especially PCBs and DDTs) (Jepson et al. 2016), reduced prey availability, habitat degradation, disturbance from vessel traffic (Lusseau et al. 2009; Williams et al. 2009; Pirotta et al 2015), entanglement and incidental bycatch, direct hunting, marine construction and anthropogenic noise (Hammond et al. 2012; Meissner et al. 2015; Pirotta et al. 2015). The determination of impacts from anthropogenic habitat degradation on coastal populations requires detailed understanding of the population structure and connectivity so that defining appropriate Management Units (MUs) or Evolutionarily
Significant Units (ESUs) can be achieved and the size of the communities/populations facing these threats estimated.

1.8 Background to the project and main aims of the study

Most research effort on bottlenose dolphins in Ireland has largely concentrated on dolphins using the Shannon estuary whose abundance has been estimated as approximately 140 individuals (Ingram & Rogan 2002; Berrow et al. 1996, 2012; Ingram & Rogan 2003; Englund et al. 2007, 2008; Rogan et al. 2015). In the early 2000s, the lower part of the estuary was designated as a SAC to ensure the protection for this population (Lower River Shannon SAC, see Fig. 1.4). However, in addition to the Shannon population, bottlenose dolphins range widely on a large part of the west coast of Ireland. Photo-identification surveys targeting these west coast animals commenced in 2001 when Ingram et al. (2001) completed surveys in Brandon Bay, (Co. Kerry), Connemara (Co. Galway), Broadhaven Bay (Co. Mayo) and McSwyne’s Bay (Co. Donegal). High number of individuals were sighted during these surveys and further genetic work suggested that these mobile animals using the coastal habitats in western Ireland belong to a genetically distinct population separate from the dolphins occupying the Shannon estuary (Mirimin et al. 2011). In 2009, an abundance estimate of 171 was derived from dedicated survey data collected during surveys of north Connemara waters (Ingram et al. 2009). This estimate represented the first attempt to assess the number of animals using a site outside of the Lower River Shannon SAC, and combined with the distribution of sightings during previous survey work and the genetic results from Mirimin et al. (2011), led to the designation of a second SAC for bottlenose dolphins on the Irish west coast in 2013 (West Connacht Coast SAC). This SAC roughly covers areas in the northern parts of Connemara, and west of the Mullet Peninsula, Co. Mayo (see Fig. 1.4). However, the number of genetic samples acquired from this area was quite low (representing less than 10% of the population), therefore it remained unclear whether even finer population structure exists among these dolphins. Moreover, the abundance surveys in 2009 were restricted to a relatively small area in Connemara, and from the comparison of photographed animals to the photo-identification archives collected over the years it was apparent that the animals were ranging well beyond this area with matches of individuals as far apart as Co. Cork and Co. Donegal. Therefore, one of the aims of this PhD thesis was to re-
investigate the population structure of bottlenose dolphins in Irish waters by analysing more genetic samples collected in a wider coastal area outside the Shannon estuary and from stranded individuals.

![Figure 1.4 Special Areas of Conservation (SACs) designated for bottlenose dolphins, *Tursiops truncatus*, in Irish waters.](image)

As detailed earlier, and especially in light of the MSFD and the recent SAC designation for the transient coastal population in Irish waters, it is essential to set up a monitoring strategy for these animals, and the first step towards an efficient monitoring plan
involves the identification of Management Units (MUs). The first and second data chapters of this thesis address this issue, first by investigating the historical divergence and colonisation patterns of the current coastal populations in the North East Atlantic (Chapter 2), whilst the recent demographic and genetic connectivity between the populations occurring in Irish waters are investigated in Chapter 3 (see Fig. 1.5).

After identifying MUs, the abundance and the scale of movements of the ‘coastal mobile’ population are evaluated in Chapter 4 using a Bayesian mark-recapture approach. Finally, on a smaller spatial scale, site occupancy and habitat use of this population are examined using passive acoustic monitoring (PAM) approaches. The effects of temporal and environmental factors influencing the bottlenose dolphin site occupancy are also evaluated in this chapter.

Figure 1.5 A simplified flow chart of the contents of the thesis, which roughly follows the steps in setting a management strategy for the coastal bottlenose dolphins in Irish waters.
All the data chapters in this thesis are written and formatted as journal manuscripts:


Author contributions: *M.N.* and A.D.F. did the laboratory work sequencing the modern samples, A.D.F. trimmed the sequences, V.B., P.G., M.T.O., provided and sequenced the ancient samples, *M.N.* analysed the data, K.K. and C.G. produced the AquaMap suitable habitat maps, M.L., A.B., N.D., W.D., V.R., F.G., E.R. and S.N.I. collected and provided the modern samples. N.W. helped with the lab work, M.L. and V.I-V. provided information on the genotypes. *M.N.* wrote the manuscript. A.D.F., E.R. and S.N.I. supervised the project and commented on the manuscript.

**Chapter 3:** M. Nykänen, E. Dillane, A. Engelund, A. D. Foote, S. N. Ingram, M. Louis, L. Mirimin, M. Oudejans and E. Rogan. *In review.* Quantifying dispersal between marine protected areas by a highly mobile species, the bottlenose dolphin, *Tursiops truncatus*

*M.N.* and A.D.F. conceptualised the work and the analyses. E.D. and *M.N.* performed laboratory work on the new samples and E.D. and L.M. did the lab work on the existing samples. *M.N.* analysed the genetic data. *M.N.*, M.O., A.E. and S.I. collected the photo-identification data, and *M.N.* analysed it. *M.N.*, S.I., E.R., A.F., M.O., and A.E. collected the genetic samples. *M.N.* wrote the manuscript. A.D.F., E.R. and S.N.I. supervised the project and commented on the manuscript.

**Chapter 4:** M. Nykänen, M. Oudejans, J. Durban, E. Rogan, A. D. Foote and S. N. Ingram. *To be submitted.* Using Bayesian inference with a multi-site mark-recapture model to estimate the abundance of a mobile population of bottlenose dolphins, *Tursiops truncatus*, on the west coast of Ireland

*M.N.*, M.O., E.R., A.D.F. and S.N.I. collected the photo-id data. *M.N.* analysed the data and wrote the manuscript. J.D. provided the code for multi-site abundance estimate. S.N.I., E.R. and A.D.F. supervised the project and commented on the manuscript.
Chapter 5: M. Nykänen, S. N. Ingram, A. D. Foote and E. Rogan. To be submitted.

Passive acoustic monitoring with C-PODs reveals patterns in site occupancy of bottlenose dolphins, *Tursiops truncatus*, on the west coast of Ireland

M.N. analysed the data and wrote the manuscript. E.R., S.N.I. and A.D.F. supervised the project and commented on the manuscript.
References


Charlton-Robb, K., Gershwin, L., Thompson, R., Austin, J., Owen, K., McKechnie, S., 2011. A New Dolphin Species, the Burrrunan Dolphin *Tursiops australis* sp. nov., Endemic to Southern Australian Coastal Waters. PLoS ONE 6, e24047. doi:10.1371/journal.pone.0024047


Please note that Chapters 2 & 3 (pp.42-138) are unavailable due to a restriction requested by the author.

CORA Cork Open Research Archive http://cora.ucc.ie
Chapter 4: Using Bayesian inference with a multi-site mark-recapture model to estimate the abundance of a mobile population of bottlenose dolphins, *Tursiops truncatus*, on the west coast of Ireland

M. NYKÄNEN, M. OUDEJANS, J. DURBAN, A. D. FOOTE, E. ROGAN and S. N. INGRAM
Abstract

Past research effort on bottlenose dolphins in Ireland has largely concentrated on a semi-resident population occupying the Shannon Estuary and the associated Lower River Shannon SAC designated for this population, but less is known about a coastal population using multiple sites throughout the Irish west coast. Especially in light of the MSFD and the recent designation of a separate SAC (West Connacht Coast SAC) for this transient coastal population, it is essential to set up a monitoring strategy for these animals. In this study, mark-recapture methods were applied to Bayesian inference and hierarchical log-linear likelihood to derive a multi-site abundance estimate of coastal bottlenose dolphins for the wider Connemara-Mayo-Donegal area in 2013 and 2014. Well-marked individuals photographed during dedicated boat-based surveys at different sites on the Irish west coast were identified from high-quality photographs taken May–August 2013 and June–September 2014 and included in the analysis. The model-averaged median estimate for the abundance of well-marked and unmarked dolphins in the whole study area in 2013 was 145 (CV = 0.30, 95% HPDI = 111–239), and in 2014 it was 189 (CV = 0.11, 95% HPDI = 162–232). High rates of movement between the different areas were observed during the study. The dolphins used the entire study area during the two years of survey effort, with nearly half (43%) of all well-marked identified animals sighted in more than one of the three survey blocks during 2013–2014. Given that the SAC designated for these animals covers a substantial area of the west coast of Ireland, the Bayesian multi-site approach is appropriate and can be applied for monitoring this population. It is well-suited for sparse recapture data collected opportunistically at multiple sites when systematic line-transect surveys are often unfeasible due to changeable weather conditions and unpredictable occurrence of the animals. With a mobile population with such extensive movements such as this one, monitoring the abundance on the dedicated SAC alone would likely lead to underestimation of the true numbers of this population.
4.1. Introduction

The size and distribution of animal populations can change in time due to variety of reasons. Continuous assessment of the status of populations and determining the reasons for these changes are key concepts of conservation research in the attempts towards preserving biodiversity. The identification of management units (MUs), usually defined as demographically independent populations whose growth rate depend almost solely on intrinsic birth and death rates rather than immigration (Palsbøll et al. 2006) is the initial step towards successful conservation and management strategies. Marine Protected Areas (MPAs) can then be designated for the species/populations to cover at least their core range, and subsequent monitoring of the abundance and others demographic parameters commenced. Obtaining accurate estimates of population size and determining trends in abundance due to demographic changes such as increased mortality or emigration should be an integral part of any management strategy allowing any changes such as a possible shift in distribution (e.g. Wilson et al. 2004, MacLeod et al. 2005) to be detected early allowing more time for appropriate management actions.

As a cosmopolitan species, bottlenose dolphins, *Tursiops truncatus*, are found in coastal inshore waters, in continental shelf regions and in open ocean environments (Wells & Scott 2002), and their minimum worldwide abundance estimate totals to about 600,000 individuals (Hammond et al. 2012) and their abundance in European waters is around 16,000 (Hammond et al. 2013). While bottlenose dolphin as a species is not considered to be globally endangered, some populations, especially the ones inhabiting coastal areas, are small and often genetically and/or geographically isolated (Natoli et al. 2005; Parsons et al. 2002, 2006; Baird et al. 2009; Rosel et al. 2009; Fernández et al. 2011; Mirimin et al. 2011; Ansmann et al. 2012; Martien et al. 2012; Caballero et al. 2012; Gaspari et al. 2015a, 2015b; Louis et al. 2014). This puts them at risk of losing heterozygosity due to genetic drift alone (Lacy 1987) and thus may affect their ability to cope with different environmental stressors. Moreover, occupying coastal habitats potentially exposes them to high levels and a wide range of human impact.

All cetaceans are listed in Annex IV of European Union’s Habitats Directive necessitating ‘strict protection’ for such species. In addition, bottlenose dolphins are
listed in Annex II of the Habitats Directive; this means that the Member States are required to designate Special Areas of Conservation (SACs) as part of European strategy to maintain or restore a favourable conservation status for the species (Natura 2000). Previous research effort on bottlenose dolphins in Ireland has largely concentrated on animals inhabiting the large open estuary of the River Shannon (e.g. Berrow et al. 1996, 2012; Ingram & Rogan 2002, 2003; Englund et al. 2007, 2008; Foley et al. 2012), and this population has been designated a SAC covering a considerable area of the lower part of the estuary (i.e. Lower River Shannon SAC). This area has been monitored since mid-1990s with summer abundance estimates calculated using data collected during several summer seasons. Estimates of abundance for dolphins using the Shannon SAC range from 114 to approximately 140 (Berrow et al. 1996, 2012; Ingram & Rogan 2002, 2003; Englund et al. 2007, 2008).

However, recent studies suggest that the bottlenose dolphins using the coastal habitats along the west coast of Ireland belong to two small, genetically and socially distinct populations which are further distinguished from a larger offshore population (see Chapter 2; Mirimin et al. 2011; Oudejans et al. 2015). Compared to the semi-resident population inhabiting the Shannon Estuary, much less is known about this second coastal population. Preliminary studies identified a significant number of bottlenose dolphins using the waters off the west coast of Ireland, and multi-annual re-sightings of individuals indicated that these animals belonged to a discrete assemblage of animals that appeared to be highly mobile with encounters occurring throughout the west coast (Ingram et al. 2001, 2003). Consequently, a second SAC for bottlenose dolphins was designated in Connemara, Co. Galway and western County Mayo (West Connacht Coast SAC) to protect this ‘coastal mobile’ population. In 2009, an abundance estimate of 171 (95% CI = 100–294) was derived from dedicated survey data collected during surveys of north Connemara waters (Ingram et al. 2009). This estimate represented the first attempt to assess the number of animals using a site outside of the Lower River Shannon SAC. However, the surveys were restricted to a relatively small area in Connemara, and from previous photo-identification work it was apparent that the animals were ranging well beyond this area with matches of individuals as far apart as Co. Cork and Co. Donegal (Ingram & Rogan 2003). Such large range and unpredictable movements make monitoring this population and deriving robust estimates of abundance especially challenging. Therefore, the aim of
this study was to obtain a more precise estimate of the number of bottlenose dolphins occupying a wider area on the west coast of Ireland, including the West Connacht Coast SAC and extending to Donegal Bay, Co. Donegal. Here, a Bayesian method suitable for opportunistic data collected from multiple sites over a wide geographic area (Durban et al. 2005) is used to estimate the abundance of this mobile population. In addition to estimating abundance, the ranging behaviour of identified individuals belonging to this transient population is examined.

4.2. Materials and methods

Boat-based photo-identification surveys

Three regional survey areas around the west and north-west coasts of Ireland, where bottlenose dolphins are frequently reported, were selected as the focus of this study. These areas broadly represent the coastal waters of north Connemara, Co. Galway, areas around Mullet Peninsula, Co. Mayo, and northern parts of Donegal Bay, Co. Donegal (Fig. 4.1), spanning over 300km along the coast. Dedicated boat-based surveys from a 6.5m rigid-hull inflatable boat (RIB) were conducted within these coastal blocks during the summers of 2013 and 2014. These surveys on predetermined tracks were conducted in Beaufort sea-states ≤3 with suitable ambient light and swell conditions with the purpose covering as much coastal area as possible in a single day, and if the weather conditions deteriorated, the survey was abandoned. In addition, opportunistic surveys (responses to sightings from the public) were conducted with the same aim as with the dedicated surveys, to locate and photograph schools of bottlenose dolphins.

A bottlenose dolphin school was defined as “all dolphins within a 100m radius of each other” (Irvine et al. 1981) and hereafter ‘encounters’ refer to periods of data collection whilst with dolphin schools. When sighted, dolphins were approached slowly and carefully, minimising changes in vessel speed and direction in order to reduce disturbance to the animals. Schools of dolphins were approached from a course that was parallel and convergent to the heading of the dolphins. Best efforts were made to photograph the dorsal fins of all members of the school during each encounter. Identification photographs were taken using a digital SLR camera (Canon EOS 1DS
mark II and Nikon D7100, 70-200mm telephoto lens) as close to perpendicular to the animals’ dorsal fin as possible and preferably within a distance of 20m.

Figure 4.1 The three coastal ‘blocks’ surveyed during 2013-2014 and the West Connacht Coast Special Area of Conservation (SACs) designated to protect this population of bottlenose dolphins. Dots denote the location of encounters with bottlenose dolphin schools during the study period.

Photograph analysis

Individual bottlenose dolphins can be identified using their natural markings (Würsig & Würsig 1977; Würsig & Jefferson 1990). These marks mostly consist of scars and nicks from interactions with conspecifics and they can be permanent, such as deep nicks or scars on the dorsal fin, or temporary, such as superficial scratches (Fig. 4.2). Other types or marks include distinctive fin shape, body deformities and lesions on the dorsal fin, flank or peduncle. Permanent marks by definition are likely to last many years, enabling long-term identification of these dolphins. In contrast, temporary markings, such as superficial tooth rakes may fade within a relatively
short period of time and inter-annual re-sighting probabilities of these animals are likely to be reduced.

**Figure 4.2** Examples of bottlenose dolphin fins showing the three grades of mark severity used in photograph analysis. Each dolphin was graded from one to three as follows: (A) grade M1 marks, consisting of significant fin damage or deep scarring that were considered permanent; (B) grade M2 marking that consist of deep tooth rakes and lesions, with only minor cuts present; (C) fin with grade M3 marks, having only superficial rakes and lesions. Grade M1 (and to some extent, M2) are considered to last many years, enabling long-term identification of these dolphins. In contrast, ‘superficial’ markings (grade M3), such as tooth rakes may fade and heal within a relatively short period of time and inter-annual re-sighting probabilities of these animals are likely to be reduced.

Digital photographs of dolphins were processed following methods described by Englund et al. (2007). For each encounter, the best quality picture was chosen of each identifiable dolphin and the quality of the photograph was graded from 1 to 4 (1 being the best) with no consideration concerning the degree of marking of the dolphin (Table 4.1). Each photographed individual was then assigned one of three grades of mark-severity (Fig. 4.2), and visually matched against the full catalogue/archive of dolphins photographed during previous encounters.

To minimise bias in photograph selection (tendency to favour photographs of well-marked animals due to their distinctiveness), the photographs were graded on their quality first before assigning marking severity grades and only the photographs of sufficient quality (Q1-3) were selected for the abundance estimation. After this, only the “well-marked” dolphins (M1) identifiable from both the left and the right side were selected in order to avoid errors in identification. Photographs from different encounters were compared within and between the regional study sites (i.e. Connemara, Mullet peninsula and Donegal Bay) separately for 2013 and 2014 to establish whether individuals were seen across the whole study area during each year. In addition, a discovery curve was fitted in order to investigate the rate at which
newly identified dolphins were added to the photo-id catalogue during the study period. Over this length of time the impact of changes in marks will be pronounced and the value of heavily marked animals emphasized.

Table 4.1 Scoring criteria for the quality of identification photographs (from Englund et al. 2007).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Well lit and focused photograph taken perpendicular to the dorsal fin at close range</td>
</tr>
<tr>
<td>Q2</td>
<td>More distant and less well lit and/or focused or slightly angled photograph of the dorsal fin</td>
</tr>
<tr>
<td>Q3</td>
<td>Poorly lit or to some extent out of focus photograph, or a photograph taken at an acute angle to the fin</td>
</tr>
<tr>
<td>Q4</td>
<td>Poorly focused, backlit or angled photograph taken at long distance to the animal</td>
</tr>
</tbody>
</table>

Multi-site abundance estimation

Mark-recapture is a widely applied numerical tool in ecology to estimate the number of identifiable individuals in a population (Otis et al. 1978). Models, such as \( \hat{M}_\text{th} \) (Chao et al. 1992), that assume population closure on a single site, are typically used in the mark-recapture abundance estimation of dolphins (e.g. Wilson et al. 1999; Read et al. 2003; Ingram et al. 2009). Even though the assumption of closure is often violated due to temporary or permanent movement of animals, it is possible to minimise this bias by decreasing the length of the sampling period. However, when animals are encountered in multiple discrete sites during a field season, these methods are no longer suitable as it is often impossible to sample throughout the whole range of the community during a single season, and this can lead to underestimation of the population size (Durban et al. 2005). Conversely, combining abundance estimates calculated separately for each site could lead to double counting of individuals due to movements between sites and thus to inflated abundance estimates due to unquantified dependency between sites.

Due to the wide and discontinuous combined survey area and sometimes opportunistic nature of the data collection (when encounters were a result of responses to sighting reports) combined with the large scale movements of the dolphins, we applied Bayesian inference to a model of hierarchical log-linear likelihood of counts of
identified dolphins across multiple discrete sites, and derived a combined abundance estimate of dolphins using the entire survey area (Connemara – Mullet peninsula – Donegal Bay) separately for the years 2013 and 2014. This method developed by Durban et al. (2005) is well-suited for sparse data sets (with low number of individual re-sightings) often associated with cetacean mark-recapture sampling and for situations when it is unfeasible to do systematic surveys covering the entire area where the animals are ranging (Durban et al. 2005). The model also takes into account the geographical dependencies between the different sites due to for example distance, thus enabling the estimation of the extent of movement between sampling locations. This approach was previously used by Cheney et al. (2013) in the abundance estimation of bottlenose dolphins around the entire Scottish coast. An advantage of using Bayesian inference instead of traditional frequentist statistics, on the other hand, is that prior knowledge of the parameter distribution can be incorporated into the model thus producing a joint posterior distribution for the parameters in question. An example of this would be setting a realistic maximum value to the prior for the unseen well-marked animals in an area. This informative prior is then incorporated into the model to facilitate the convergence of Markov Chain Monte Carlo (MCMC) chains.

A contingency table of the sightings histories of identified marked bottlenose dolphins was created based on their presence or absence in each of the study blocks during each season (see Table 4.2) and implemented in the model. The resighting of individuals in multiple survey sites thus represents “capture-recapture events” defined by space instead of time in the context of traditional capture-recapture abundance estimation (Durban et al. 2005). The missing value (NA) on the last row of Table 4.2 represents the number of individuals that were not seen in any of the study blocks (missed well-marked dolphins) during the season, and the purpose of the model is to predict a value for the missing cell and thus estimate the overall abundance of well-marked animals across all of the study sites. Specifically, the cell counts (Table 4.2) are treated as independent Poisson random variables with a mean \( \mu_i \), so the logarithm of the Poisson mean can be modelled as additive regression function of study area effects:

\[
\log(\mu_i) = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1} x_{i2} + \beta_5 x_{i1} x_{i3} + \beta_6 x_{i2} x_{i3}
\]

where \( \beta \) is the vector of unknown parameters to be estimated and the \( x_i \)'s are indicator variables for the study block classifying factors in the design of the contingency table (Durban et al. 2005). For example, \( x_{i0} = 1 \) for all \( i \), and \( \beta_0 \) thus equals to an overall
mean of the counts on the logarithmic scale. The indicators $x_{i1}$, $x_{i2}$, $x_{i3}$ then take values of either 1 or -1 depending on the attribute 1 = seen, or -1 = not seen, for study blocks 1, 2, and 3 (or Connemara, Mullet Peninsula and Donegal Bay), respectively (see Table 4.2). The parameters $\beta_1$, $\beta_2$, $\beta_3$ thus denote the main effects of each of the study areas on the overall mean $\beta_0$, describing the difference between the average of the $\mu_i$'s for cells relating to each study block, and the average of all $\mu_i$'s. The terms containing products of any two of these indicators describe two-way interaction effects, with $\beta_4$, $\beta_5$, and $\beta_6$ reflecting the strength and direction of movements of animals from pairs of study sites 1:2, 1:3, and 2:3 respectively. Different models can be produced by omission of one or more of these interaction effects, and so a model averaged estimate for the total number of well-marked individuals, weighted by the relative probability of the candidate models, was produced. This model averaging and prediction were performed using MCMC sampling in WINBUGS software (Lunn et al. 2000) with 100,000 burn-in followed by 100,000 iterations. Three chains were run in order to confirm consistency between runs and inspected visually for convergence. The model also incorporates the proportion of well-marked individuals as a binomial sample of the total number of all animals seen (regardless of marking severity); therefore it predicts the number of all individuals in the study area (see Cheney et al. 2013).
Table 4.2 Contingency table showing the count of well-marked (M1) bottlenose dolphins present (1) or absent (-1) in each of the study sites Connemara, Mullet peninsula and Donegal Bay in 2013 and 2014.

<table>
<thead>
<tr>
<th>Year 2013</th>
<th>Count</th>
<th>Connemara</th>
<th>Mullet peninsula</th>
<th>Donegal Bay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year 2014</th>
<th>Count</th>
<th>Connemara</th>
<th>Mullet peninsula</th>
<th>Donegal Bay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>

As a further attempt to describe the ranging behavior of bottlenose dolphins sighted on the west coast of Ireland, the range of sighting latitudes were plotted of the 75 most sighted (≥5 times) well-marked dolphins encountered since the photo-identification catalogue for these animals was started in 2001.

Program TRENDS (Gerrodette 1993) was used to conduct a power analysis in order to assess the amount of sampling effort (in this case, number of years) required to detect a yearly decline of 10% in population size using the coefficients of variation (CV) obtained for the derived abundance estimates. In addition, the effect of different amount of sampling effort on the minimum detectable overall decline in population size was examined with CVs varying from 0.01 to 0.30. Specifically, scenarios were tested when the population was sampled once, every three years, every two years or every year, during a period of six years (which is the reporting period set in the Habitats
Directive). In all the power analyses, the probability of Type I and II errors was set to 0.05, and a one-tailed test was used, as the purpose was to detect a decrease and not a general change in abundance. The CV was chosen as proportional to the square root of abundance (as recommended by Gerrodette 1987 for mark-recapture sampling studies) and an exponential population model used as per Daura-Jorge et al. (2013).

4.3. Results

Survey effort and encountered bottlenose dolphin schools

Survey effort varied between years and sites (Table 4.3) with total number of 174 survey hours in 2013 and 146 survey hours in 2014. The number of encounters also varied between sites and years, with eight bottlenose dolphin schools encountered in Connemara, three in the waters around Mullet peninsula and two in Donegal Bay in 2013. In contrast, six dolphin schools were encountered in Connemara, seven around the Mullet peninsula, eight in Killala Bay/south Donegal in 2014 (Fig. 4.3). Bottlenose dolphin school size varied between encounters, locations and years, with larger median school sizes observed in 2014 (Table 4.4). The largest schools were encountered in Donegal Bay with median sizes of 39 and 36 in 2013 and 2014, respectively.

Abundance and movements

Abundance estimates were calculated for each year separately. A total of 59 well-marked bottlenose dolphins were selected for the analysis as identified from the high-quality photographs taken in May–August 2013 (Table 4.1). Fifteen of these individuals (25%) were recorded in more than one of the study areas, with similar numerical overlap between Connemara and Mullet peninsula (six dolphins), Mullet peninsula and Donegal Bay (two dolphins) and Connemara and Donegal Bay (four dolphins). Three out of the 59 well-marked dolphins (5%) were seen in all of the study blocks (Table 4.2). Posterior model probabilities with site interaction terms are presented in Table 4.5., and the model averaged median estimate for the abundance for the whole study area in 2013 was 145 (CV = 0.30, 95% HPDI = 111–239). This estimate includes all well-marked, less marked and unmarked individuals (excluding calves) thus representing the total abundance of dolphins.
Table 4.3 Yearly survey effort in bottlenose dolphin surveys 2013-2014.

<table>
<thead>
<tr>
<th></th>
<th>Survey effort (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2013</td>
</tr>
<tr>
<td>Connemara</td>
<td>118</td>
</tr>
<tr>
<td>Mullet peninsula</td>
<td>29</td>
</tr>
<tr>
<td>Donegal Bay</td>
<td>27</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>174</strong></td>
</tr>
</tbody>
</table>

From the photo-identifications from May–September 2014, a total of 91 well-marked dolphins were included in the analysis (Table 4.2). Eight dolphins (9%) were encountered in all of the study areas. The highest overlap in site use was between Mullet peninsula and Donegal Bay with 28 dolphins (31%) sighted in both these areas. Donegal also had the highest number (23 individuals) of animals seen in only one of the three study sites. The Bayesian multi-site abundance median estimate of the total number of dolphins for the whole study area for the summer 2014 was 189 (CV = 0.11, 95% HPDI = 162–232).

Table 4.4 Median school sizes (with minimum and maximum) of bottlenose dolphins encountered during abundance surveys in 2013-2014.

<table>
<thead>
<tr>
<th></th>
<th>School size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2013</td>
</tr>
<tr>
<td>Connemara</td>
<td>8.5 (4 - 25)</td>
</tr>
<tr>
<td>Mullet peninsula</td>
<td>17 (10 - 31)</td>
</tr>
<tr>
<td>Donegal Bay</td>
<td>39 (30 - 48)</td>
</tr>
<tr>
<td><strong>All areas</strong></td>
<td><strong>12 (4 - 48)</strong></td>
</tr>
</tbody>
</table>

The number of well-marked (M1) individuals photographed on the west coast of Ireland between 2013 and 2014 is presented as a discovery curve in Fig. 4.3. The gradual levelling off of the rate of discovery of previously uncatalogued animals could indicate that most individuals occurring within the area were photographed and archived, with fewer new individuals being added towards the end of the sampling period.
Table 4.5 Posterior model probabilities, $P$, of eight log-linear candidate models corresponding to the inclusion of different sets of interaction terms ($\beta_i x_{ii}$) between study sites Connemara, Mullet Peninsula and Donegal Bay. The averaged model was then used to derive abundance estimates for bottlenose dolphins for 2013 and 2014 across all the models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model terms</th>
<th>$P$ (model)</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>No interaction – only main effects</td>
<td>$\beta_1 x_{i1}, \beta_2 x_{i2}, \beta_3 x_{i3}$</td>
<td>0.07</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Connemara:Mullet peninsula</td>
<td>$\beta_4 x_{i1} x_{i2}$</td>
<td>0.52</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Connemara:Donegal</td>
<td>$\beta_5 x_{i1} x_{i3}$</td>
<td>0.02</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Mullet peninsula:Donegal</td>
<td>$\beta_6 x_{i2} x_{i3}$</td>
<td>0.16</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Connemara:Mullet peninsula,</td>
<td>$\beta_4 x_{i1} x_{i2}, \beta_5 x_{i1} x_{i3}$</td>
<td>0.03</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Connemara:Donegal</td>
<td>$\beta_4 x_{i1} x_{i2}, \beta_6 x_{i2} x_{i3}$</td>
<td>0.13</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Connemara:Mullet peninsula,</td>
<td>$\beta_5 x_{i1} x_{i3}, \beta_6 x_{i2} x_{i3}$</td>
<td>0.01</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Connemara:Donegal, Mullet</td>
<td>$\beta_4 x_{i1} x_{i2}, \beta_5 x_{i1} x_{i3}, \beta_6 x_{i2} x_{i3}$</td>
<td>0.05</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Connemara:Donegal, Mullet</td>
<td>Averaged</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.3 Discovery curve for bottlenose dolphin identifications made during photo-identification surveys in 2013–2014. Only well-marked (M1) dolphins identifiable from both sides are included in the data.
The range of sighting latitudes of the most sighted well-marked dolphins identifiable from both sides is presented in Fig. 4.4; it appears that while most of these animals were sighted from Donegal Bay to Connemara with the distance between the areas of nearly 250km, there were some animals that had even wider distribution having been sighted from Co. Cork to Donegal Bay between 2001 and 2014 with over 500km between these sites. In contrast, there were also a few individuals that were only seen in the Connemara study block.

The results from power analysis showed that with a coefficient of variation (CV) of 0.11, the uncertainty associated with the 2014 abundance estimate, it would take six years of annual survey effort to detect a trend of 10% annual decline in abundance with 95% probability; this equates to 47% overall population decline. Moreover, with the larger uncertainty around the 2013 estimate (CV = 0.30), it would require 9 years of annual surveys to detect the same annual trend; this would total to 61% overall decline. The effects of different sampling schemes and CVs on the minimum detectable change in population size are presented in Fig. 4.5. If abundance was sampled twice, as in this study but assuming uniform sampling periods instead of sampling on consecutive years, the minimum detectable overall change in a population during a six year monitoring period would be 29% with a CV of 0.11 (Fig. 3). With the CV of 0.30, however, the population would have to decline to less than 50% of its original level before it could be detected with 95% certainty (Fig. 4.5).
Figure 4.4 The geographic range of the 75 most sighted individually identified bottlenose dolphins. The outline of Ireland is given only for reference and has been scaled to correspond the sighting latitudes. Data for the figure were collected 2001-2014 with only individuals sighted at least five times included. The center line and the bottom and top of the box represent the 50th, 25th and 75th percentiles, respectively, and the whiskers the 5th and 95th percentile. The dots represent “outliers” in the data. The data have been ordered with increasing median latitude.

Figure 4.5 The effect of coefficient of variation (CV) to the minimum detectable change in the abundance of a theoretical population with different levels of sampling effort spread uniformly during a 6-year period.
4.4. Discussion

During 2013–2014 survey seasons 83 new dolphin identifications from photographs obtained from both sides of the animal were added to the catalogue. In addition, 38 animals were matched to animals identified from encounters made in previous years with ten identifications dating back as far as 2001 (Ingram et al. 2001). Such long term re-identifications indicate that a degree of site fidelity is evident in animals using coastal waters in the west and north-west of Ireland, and it appears that together the combined area between Connemara, the Mullet peninsula and Donegal Bay form an important part of the home-range for a large number of bottlenose dolphins. Some of the dolphins used the entire study area during the two years of survey effort, with almost half of all well-marked identified animals encountered in more than one of the three survey blocks during 2013–2014. This wide scale habitat use pattern produces patchy temporal site occupancy as individuals and schools range freely over considerable distances around the Irish coast and possibly further afield.

Abundance of bottlenose dolphins on the west coast of Ireland

Of the two years of the study, summer estimates of abundance for 2014 were larger and more precise than for 2013 with an overall median multi-site estimate of 189 (CV = 0.11, 95% HPDI = 162–232) compared to 145 (CV = 0.30, 95% HPDI = 111–239), respectively. This difference in the yearly estimates, albeit not significant, may be caused by a combination of environmental, sampling and/or behavioural factors. The total number of surveys, encounters and identifications were higher during 2014 leading to a more comprehensive dataset, and the weather was notably better during that summer than in the previous year. However, the difference may also be attributable to unknown differences in the ranging behaviour of many of the dolphins in the areas surveyed during the two years due to heterogeneous and unpredictable changes in seasonal site use or possibly also due to variation in the composition of the population occupying Irish coastal waters from one summer to the next. Overall, the multi-site abundance estimates, especially the one for the summer of 2014, are comparable with a previous abundance estimate of 171 (95% CI = 100–294) for Connemara for 2009 (Ingram et al. 2009) and the cumulative number of animals (N = 179) identified in Mayo in 2008-2009 (Oudejans et al. 2010). However, even though the 2014 estimate can be considered reliable, it is likely that the actual number of
animals using the coastal waters of western Ireland falls to the higher side of the confidence interval based on a cumulative number of identifications collected by Ingram et al. (2001, 2003, 2009) and M. Oudejans who keeps a separate catalogue of the dolphins occurring in the waters around the Mullet peninsula (M. Oudejans, personal communication).

Abundance of bottlenose dolphins in other coastal areas of the Northeast Atlantic region

With the abundance estimate for dolphins using the Shannon Estuary SAC ranging from 114 to 140 (Berrow et al. 1996, 2012; Ingram & Rogan 2002; Ingram & Rogan 2003; Englund et al. 2007, 2008), the multi-site abundance estimates for the west coast of Ireland are similar to the ones for the East coast of Scotland obtained using the same multi-site approach (195 with 95% HPDI of 162–253 in 2006, and 227 with 95% HPDI of 175–384 in 2007) (Cheney et al. 2013). Another adjacent community of bottlenose dolphins are found in the Sound of Barra, in the Outer Hebrides, Scotland, but this community is significantly smaller consisting of only 6–15 individuals (Grellier & Wilson 2003). This group of dolphins is also thought to have high site fidelity with repeated identifications of the same individuals within and between years. Moreover, Cheney et al. (2013) calculated the combined abundance for the entire Scottish Western Coast, including the Sound of Barra, as 45 individuals (95% HPDI of 31–71 in 2006, and 33–66 in 2007). Yet another neighbouring semi-resident group of bottlenose dolphins is found in Cardigan Bay, Wales, with mean summer abundance estimates of dolphins using the SAC varying between 70 and 214 in 2003-2007 (Ugarte & Evans 2006; Pesante et al. 2008; Veneruso & Evans 2012; Feingold & Evans 2013). Similar to the ‘mobile’ population on the West Coast of Ireland, these dolphins are known to occupy a wider range of habitats whilst having a seasonal occupancy in Cardigan Bay during the summer months and majority of the animals moving northwards to the Irish Sea during winter (Baines et al. 2002; Pesante et al, 2008). Interestingly, based on evidence from a recent study by Louis et al. (2014), all of the above mentioned bottlenose dolphins (including the ones on the west coast of Ireland) may belong to a wider ‘Coastal North’ population thus retaining significant gene flow between the different communities. However, the ‘Coastal North’ population differs genetically from the much larger but also somewhat nearby ‘Coastal South’ population occupying the Normano-Breton Gulf of the English Channel whose
abundance was estimated as 420 dolphins (95% CI = 331–521) in 2010 (Louis et al. 2015), making this the largest community of coastal bottlenose dolphins in central-northern Europe.

**Multi-site model**

The multi-site estimates derived in this study are likely to better reflect the true abundance of the coastal bottlenose dolphins on the west coast of Ireland than previous local site-based estimates due to the wider-scale sampling over a larger coastal area. More widespread sampling effort increases the probability of encountering more of these mobile animals and accounts for the pseudoreplication of individuals using multiple surveyed sites if abundance estimates were to be derived separately for each study site. Furthermore, individual heterogeneity in capture probabilities is estimated by geographical dependencies between study sites, included in the model matrix and thus incorporated in the abundance estimate (Durban et al. 2005). However, like the Mth model by Chao et al. (1992), a common method used in cetacean abundance studies (e.g. Bearzi et al. 2008; Vermeulen & Cammareri 2009; Gnone et al. 2011; Brown et al. 2014), the Bayesian multi-site approach assumes population closure with no births, deaths, immigration or emigration occurring in the area during the study period (Durban et al. 2005). It is likely that although this assumption may be susceptible to violation due to the large scale of the animals’ ranges, the inclusion of multiple sites over a broad geographical area should improve this model’s performance compared to closed population Maximum Likelihood derived abundance estimates such as the Mth model. Furthermore, the short duration of the annual survey season (May–August in 2013 and May–September in 2014) likely reduces the rates of immigration, emigration and deaths of individuals sampled thus increasing the likelihood of effective closure of the sampled population.

**Range and movements of individuals**

Some of the bottlenose dolphins that were encountered during surveys in 2013–2014 had previously been recorded as far south as Co. Cork and appear to range widely around the west coast of Ireland and possibly beyond (Fig. 4.4). A case in point, a dolphin encountered in Donegal Bay in the summer of 2014 has previously been photographed in the Moray Firth in 2001 and around the Hebrides in 2004 (Robinson et al. 2012) thus providing further evidence of the long distance movements and
transient behaviour of at least some of these animals. The dolphins used the entire study area covered during the two years of survey effort, with nearly half (43%) of all well-marked identified animals sighted in more than one of the three survey blocks during 2013–2014. Similarly, Cheney et al. (2013) found a large percentage of dolphins (57%) using more than one study site, however, distances between the sites on the east coast of Scotland were shorter compared to the present study with over 300km between Connemara and Donegal Bay.

**Management implications**

Bottlenose dolphin populations using coastal environments are at particular risk of exposure to a number of anthropogenic threats which may directly impact individuals, for example through disturbance or damage to health and to the overall functioning of the coastal ecosystems upon which they depend. The sensitivity of bottlenose dolphins to some of these threats is exacerbated by their position as an apex predator (Jepson et al. 2016) and also by their low reproductive rates (Connor et al. 2000; Quick et al. 2014). The main threats in coastal environments include pollutants such as xenobiotic chemicals (especially PCBs and DDTs) (Jepson et al. 2016), reduced prey availability, habitat degradation, disturbance from vessel traffic (Lusseau et al. 2009; Williams et al. 2009; Pirotta et al. 2015), entanglement and incidental bycatch, direct hunting, marine construction and anthropogenic noise (Hammond et al. 2012; Williams et al. 2014; Pirotta et al. 2015). The determination of impacts from anthropogenic habitat degradation on coastal populations requires detailed understanding of the population structure and size of the communities/populations affected, and the ranging behaviour and site fidelity of individuals within these populations. The fact that coastal bottlenose dolphin populations often display fine-scale genetic structuring (Natoli et al. 2005; Parsons et al. 2002, 2006; Baird et al. 2009; Rosel et al. 2009; Fernández et al. 2011; Martien et al. 2011; Mirimin et al. 2011; Caballero et al. 2012; Gaspari et al. 2013, 2015; Louis et al. 2014), even in adjacent coastal areas (see Chapter 3; Mirimin et al. 2011) where there are no obvious physical barriers preventing gene flow, presents an added challenge to effective conservation and management since delineation of MUs is required and the amount of gene flow between them needs to be quantified before monitoring can commence.

Efficient and regular monitoring of abundance is vital to the management of coastal SACs that have been designated for bottlenose dolphins. Whereas other SACs
designated for the protection of this species around Ireland and the UK appear to describe a considerable degree of site fidelity in a single confined bay (e.g. Shannon estuary, Sound of Barra, Cardigan Bay, Moray Firth) this mobile west/north-west coast population does not fit this pattern and thus represents challenges to designing an effective and robust monitoring strategy. Such a strategy must provide accurate data that will contribute to knowledge of the status of the conservation feature and also enable the competent authority to detect, in a timely manner, changes in abundance, population viability and survival rates in order to assess conservation status. The Bayesian multi-site approach used in this study appears to provide a precise and comprehensive estimate of the abundance of dolphins in this wider habitat area and is useful in informing the management of the SAC. With a mobile population with extensive movements such as this one, monitoring the abundance on the dedicated SAC alone would likely lead to underestimation of the true numbers of this population.

In order to be able to detect an overall decline of 25% in abundance over a six-year reporting period, a guideline set in the Article 17 of the Habitats Directive for a population to remain at a “favourable level”, the CV around the abundance estimate would have to be as low as 0.08. This would be difficult, if not impossible, to achieve for a mobile population such as the one inhabiting the west/northwest coast of Ireland. An alternative strategy, where the 25% decline could be detected, would be to sample the abundance every two years. However, this would only be achievable if the CV did not exceed 0.11 between sampling occasions. In addition, it was determined that six consecutive years of annual monitoring would be required to detect a 10% annual decline in population, even with a low CV of 0.11. Thus it seems unrealistic to achieve an uncertainty small enough (i.e., CV of 0.02) that an annual decline of 1% (another guideline in Article 17) could be detected. Therefore, regular monitoring and wide-scale research effort where the population is sampled annually or at least every two years is recommended in order to be able to detect changes in population dynamics within this mobile coastal population. This will increase the ability to implement the best possible conservation strategies in a timely manner ensuring the long-term viability of this population.

According to Durban et al. (2005) the Bayesian multi-site model of abundance takes into account uncertainty from having a sparse data set and also the uncertainty in model selection by weighing the different model probabilities and thus producing a model-
averaged estimate. It also accounts for much of the individual variation in capture probabilities caused by varying extent of movement of individuals between surveyed sites. However, further heterogeneity unrelated to movement patterns, such as differential reaction of individual dolphins towards the research boat, may still exist that is not captured by the model (Durban et al. 2005) so caution should be taken when interpreting the estimates. Nevertheless, the multi-site approach seems to produce abundance estimates with less uncertainty around the point estimate. In addition, multiple discrete locations can be sampled simultaneously, and photo-ID surveys can be done opportunistically with the help from a sightings network prompting the researchers to the location of the dolphins. Monitoring and deriving abundance estimates for a single site is liable to miss animals that range widely and are not present during the period of surveys resulting in negatively biased estimates due to heterogeneity in movements of animals rather than actual changes in population size.

In order to gain a more complete picture of the scale of movements and occupancy of this population the work reported should be extended around larger sections of the Irish coast. Such work would ensure that other sites of importance to these animals are identified along with any potentially harmful interactions with human activities.

Although the full extent of the ranges of individuals in this population are not yet known, previous research has shown that at least some of these animals travel distances at the scale of hundreds of kilometres (Ingram et al. 2001, 2003; O’Brien et al. 2009; Robinson et al. 2012; Cheney et al. 2013). It seems that these animals may form part of the ‘Coastal North’ meta-population defined by Louis et al. (2014), and trans-national movements of many more individual dolphins than has been reported up to now are likely. High levels of mobility, in turn, can result in substantial gene flow and the homogenization of genetic diversity across a geographic range (e.g. Winkelmann et al. 2013), and it may be that the transient bottlenose dolphins ranging from the Irish west coast to Moray Firth should be considered as a single management unit and managed in co-operation with Ireland and the UK. As mentioned previously, a number of individuals from the west coast of Ireland have been matched on an ad-hoc basis to other existing catalogues but there is a need for a collaborative effort to consistently and regularly compare photo-id catalogues from separate regions/countries (e.g. Wales, Scotland, France, Cornwall) in order to better elucidate ranging patterns, demographic dispersal and the abundance of this putative meta-population. In
addition, genetic dispersal within the meta-population needs to be quantified through increased sampling effort over a larger area extending beyond country boundaries and using a common set of genetic markers that are comparable between laboratories. Thus delineating populations remains essential in the management of these dolphins. This is achievable via wider genetic sampling along the entire range of bottlenose dolphins occurring in the coastal areas of Ireland, UK and northern France and with use of a common set of molecular markers. Only after this delineation of MUs can a comprehensive management plans for protecting the populations be drawn.
References


Discrete or not so discrete: Long distance movements by coastal bottlenose dolphins in UK and Irish waters. Journal of Cetacean Research and Management.


Chapter 5: Passive acoustic monitoring of site occupancy by bottlenose dolphins, *Tursiops truncatus*, on the west coast of Ireland

M. NYKÄNEN, S. N. INGRAM, A. D. FOOTE and E. ROGAN
Abstract

While visual surveys are limited to good weather conditions and visibility, passive acoustic monitoring (PAM) provides a non-invasive method for continuous monitoring of vocalising species. PAM can be used to monitor habitat use, migratory patterns, abundance and foraging behaviour of cetacean and non-cetacean species. The heterogeneous and wide ranging movement patterns of ‘mobile’ bottlenose dolphins inhabiting the coastal waters of Ireland presents challenges in monitoring. In the present study, temporal and seasonal trends in dolphin occupancy were investigated at two key sites on the west coast of Ireland using dolphin echolocation detection with C-PODs and by applying Generalized Additive Models (GAMs). In addition, the effects of hydrological and environmental parameters that may influence prey distribution were examined in relation to temporal patterns in click detections. Autocorrelation in the data was modelled using Generalized Estimating Equations (GEEs). Tidal current speed and direction were found to be significantly associated with the presence of dolphins with more echolocation click trains logged at intermediate and high current speeds and in northerly and southerly running currents. It may be that the increased detections of dolphins associated with faster current speed reflects the movements of prey species. Significantly more detections were logged by the C-POD deployed at the mouth of Killary Fjord, Co. Galway than by the C-POD in McSwyne’s Bay, Co. Donegal. It may be that hydrographic features such as steep and narrow channels in the bottom of the fjord combined with currents resulting from tidal flow may gather and concentrate prey thus facilitating capture in this site. In addition to tidal parameters, an increase in dolphin echolocation click detections in the spring was also recorded in Killary Fjord. This coincides with the peak run of adult salmon, and it is possible that this seasonal increase in prey availability is a driving force for dolphin presence. This study demonstrates the potential of using PAM to reveal information on seasonal and temporal habitat use of dolphins and provides evidence that C-PODs can be used as a long-term monitoring tool in a relatively cost effective way that could potentially be implemented as a part of the monitoring requirements under the EU Habitats and Marine Strategy Framework Directives.
5.1. Introduction

Understanding animal movements, habitat use and their drivers is challenging, especially in wide ranging populations. Environmental and biological processes can influence the relative abundance and distribution of marine organisms on spatial and temporal scales. Marine mammal habitat use can be influenced by a range of drivers including prey availability and foraging strategies (Ford et al. 1998; Saulitis et al. 2000; Heithaus & Dill 2002; Sveegard et al. 2012), anthropogenic disturbance (Johnston et al. 2002; Olesiuk et al. 2002; Pirotta et al. 2014a) and predator avoidance (Heithaus & Dill 2002). Habitat use has also been linked to hydrographic and environmental parameters such as depth, slope, temperature, tidal range and current speed (Wilson et al. 1997; Ingram & Rogan 2002; Hastie et al. 2003; Hastie et al. 2004; Embling et al. 2010; Booth et al. 2013). Significant drivers of the wide-scale distribution of cetaceans, on the other hand, include at least temperature, primary productivity and ocean fronts (Jaquet & Whitehead 1996; Littaye et al. 2004; Kaschner et al. 2006; Foote et al. 2013; Scales et al. 2014, see also Fig. 2.2 in Chapter 2), factors that are susceptible to considerable temporal fluctuations. Understanding temporal and spatial habitat use of marine mammals is often challenging due to the dynamic environment that they live in.

While visual surveys are limited to good weather conditions and visibility, passive acoustic monitoring (PAM) provides a non-invasive method for continuous monitoring of vocalising species that can be elusive or otherwise visually difficult to detect. PAM is commonly recommended for mitigation of impacts on cetaceans during marine construction such as pile driving or activities related to seismic surveys (JNCC 2010; DAHG 2014). PAM can also be used in studies on habitat use, migratory patterns, abundance and even foraging behaviour and depredation of cetacean and non-cetacean species (Norris et al. 1999; McDonald & Fox 1999; Carlström 2005; Verfuß et al. 2007; Luczkovich et al. 2008; VanParijs et al. 2009; Rountree et al. 2011; Kyhn et al. 2012; Marques et al. 2013; Pirotta et al. 2014b; Thode et al. 2015; Benjamins et al. 2016).

Types of acoustic monitoring devices used to detect cetaceans vary from extensive static hydrophone arrays such as SOSUS and AUTEC (deployed by the US Navy for submarine surveillance) to single stationary hydrophones and towed hydrophone
arrays (Mellinger et al. 2007). These systems usually record sounds on a certain frequency band and can be selected based on the vocalisation range of the target species. They can either collect and store data continuously (like the SOSUS array and various self-contained units such as the archival marine acoustic recording unit (ARU), Ecological Acoustic Recorder (EAR) and Wildlife Computers (SM2M), or they can be set to record for a period of time on a predefined interval. The recordings are either stored in the unit itself which can then be retrieved, or the hydrophone can be linked via a cable to a listening station allowing for real-time analysis of the sound data. Unlike conventional hydrophones, C-PODs (or T-PODs, older versions of C-PODs) are commonly used for detecting and monitoring odontocetes particularly in coastal habitat (Chelonia Ltd, Cornwall, UK). C-PODs are self-contained battery-powered acoustic devices that detect echolocation clicks or bouts produced underwater by toothed whales like dolphins and porpoises mainly for orientation and localization in foraging (see Tyack and Miller 2002). The conservative detection radius of C-PODs has been estimated to be ~300–400m for bottlenose dolphins depending on the behaviour of the animals, but detections up to over 1,500m have been recorded (Nuuttila et al. 2013). These autonomous loggers, which are capable of collecting continuous time-series data for several months between service intervals, do not record sound but instead log the time, duration, inter-click interval, dominant frequency and other features of detected click trains (defined as consisting of at least five consecutive clicks) with up to a 5ms resolution. The click bouts detected and logged by the C-POD are saved onto an SD card, from which the data can be downloaded and analysed. However, even though the filtering algorithms that come with the C-POD software can discriminate between narrow-band high-frequency clicks produced by porpoises and broad-band dolphin echolocation clicks, the identification of individual animals or the effective discrimination of clicks from different dolphin species is not currently possible (Tregenza 2013) due to the high overlap in click train characteristics (Robbins et al. 2015). Nevertheless, C-PODs (and T-PODs) provide a reasonably economic way to collect information on the site use of echolocating odontocetes.

The common bottlenose dolphin, *Tursiops truncatus*, has been allocated a special protected status under Annex II and Annex IV of the European Union’s Habitat’s Directive, which means that the status of the species must be maintained at or restored to a ‘favourable condition’. To protect critical areas for the two genetically and socially
distinct coastal populations inhabiting Irish coastal waters (Mirimin et al. 2011; Chapters 3 and 4), two Special Areas of Conservation (SACs) for bottlenose dolphins have been designated, one in the Shannon estuary (Lower River Shannon SAC) and one in two areas in Connemara and Mayo (West Connacht Coast SAC) (Fig. 5.1).

The wide geographic ranging patterns of members of the ‘mobile’ west coast population, estimated to comprise 189 dolphins (median estimate, with a CV of 0.11, see Chapter 3 on abundance), presents significant monitoring challenges. Therefore, in the present study, temporal and seasonal trends in dolphin occupancy were investigated at two key sites on the west coast of Ireland using dolphin echolocation detection with C-PODs. In addition, the effects of hydrological and environmental parameters such as tide height, speed and direction, primary productivity and Sea Surface Temperature (SST), were examined in relation to temporal patterns in click detections and dolphin site occupancy, and the potential of using PAM as a monitoring tool in the management of marine protected areas is discussed.

5.2. Materials and Methods

C-POD deployment and study sites

The two deployment locations in Western Ireland were selected, one at the mouth of Killary Fjord, Co. Galway (part of the West Connacht Coast SAC), and the other in McSwyne’s Bay, Co. Donegal (Fig. 5.1.), based on prior knowledge of bottlenose dolphin habitat use (Ingram et al. 2001, 2009). Killary Fjord is one of the three glacial fjords situated in Ireland. The fjord is approximately 15km long and 0.75km wide with an average depth of 15m and a maximum depth of 45m. Small scale commercial draft netting fisheries operate in the fjord with an annual catch of 369 salmon reported in 2012 (Anonymous 2012), and the three tributaries feeding into the fjord are important rivers for recreational salmon fishing with the combined harvest of 287 fish in these rivers in 2012 (Anonymous 2012). In addition, Killary Fjord is an important site for shellfish aquaculture, and a salmon farm is situated immediately outside the mouth of the fjord. McSwyne’s Bay, on the other hand, is approximately 8km wide with an average depth of 20-35m. Similar to Killary Fjord, McSwyne’s Bay is also an important area for aquaculture with salmon and mussel farms in proximity of the C-POD deployment site.
The C-PODs used in this study were deployed very close to shore to minimise the detection of other delphinid species (e.g. Perrin 2002; Bearzi et al. 2003), in waters of 6-12m depth (at low tide) and 2-3m of the bottom of the seafloor. The C-POD in McSwyne’s Bay, Co. Donegal, was deployed on the 19th of October 2013, and a second C-POD was deployed in Killary Fjord, Co. Galway, on the 15th of September 2014 following a theft of a C-POD previously deployed in early June 2014. The C-PODs were retrieved every 4-6 months for data download and battery replacement (subject to weather conditions) and were re-deployed as soon as possible after the maintenance.

**Figure 5.1** Locations of C-PODs deployed to detect bottlenose dolphins using the West Connacht Coast SAC and a previously identified ‘hotspot’ in Donegal Bay (Ingram et al. 2001).

C-POD data were analysed in C-POD.EXE software (Chelonia Ltd.) using the GENENC click classifier and ‘other cet’ setting which maximises the capture of echolocation click events for dolphins (N. Tregenza, Chelonia Limited, personal communication). The classifier also discriminates broadband dolphin clicks from narrow-band high frequency harbour porpoise (*Phocoena phocoena*) clicks, boat engine noise and background environmental noise such as sounds caused by the moving sediment.
Environmental data

The environmental variables were chosen for this study partly due to their significant effect on dolphin presence found in a previous study (i.e. tidal parameters, Pirotta et al. 2014) or due to their potential effects on prey availability driving the distribution of marine predators (Mendes et al. 2002). Remotely sensed surface chlorophyll-a data from the MODIS sensor on the NASA’s Aqua satellite were downloaded as monthly values averaged over 4km × 4km grid cells within an equivalent area surrounding each of the two deployment sites. These data were downloaded using the GIOVANNI portal on the NASA website (http://giovanni.gsfc.nasa.gov/giovanni/). Real time hourly measurements of Sea Surface Temperature (SST) in Galway Bay were downloaded for the duration of the deployment from the Marine Institute website (http://www.marine.ie/Home/site-area/data-services/real-time-observations/wave-buoys) and used as a large scale proxy for SST at both deployment sites. Time-series data on sunrise and sunset times for the west coast of Ireland (54.451°N, 9.297°W) were retrieved using R (R Core Team 2016) package ‘maptools’ (Bivand and Lewin-Koh 2016); the functions use algorithms provided by the National Oceanic & Atmospheric Administration (NOAA). Hourly tidal level, tidal current speed and direction were obtained using POLPRED (NERC National Oceanography Centre, Liverpool, UK). However, predictions of tidal information at locations immediately close to the coast are not available in POLPRED; therefore the predictions were made to the closest possible grid cell to the C-POD deployment locations, within 8km from both locations.

Statistical modelling: GEE-GAMs on time-series data

The R (R Core Team 2016) package ‘MRSea’ (Hayward et al. 2013) with Spatially Adaptive Local Smoothing Algorithm (SALSA, Walker et al. 2010) was used to fit splines to the continuous covariates in generalized additive models (GAMs; Wood 2006). SALSA is an automated procedure that finds the best way to fit a regression spline for one or two-dimensional covariates and performs knot selection, or otherwise reduces the covariate to a linear term (Walker et al. 2010; Hayward et al. 2013). Due to different temporal resolution in the environmental data sets, explanatory variables included in the GAMs were divided into three different models: an hourly model, a daily model and a monthly model. The hourly model included a factor covariate ‘daylight’ with two levels ‘light’ and ‘dark’, and continuous tidal covariates ‘tidal...
level’, ‘tidal current speed’ and ‘current direction’; in general, this model included
covariates measured on an hourly scale. The response variable was modelled as
presence/absence (of dolphin echolocation click trains per hour) with a binomial
distribution due to the fact that the model residuals were not over- or under-dispersed.
The daily model included the factor variable ‘site’ (deployment location) and temporal
covariates ‘year’ and ‘Julian day’, the latter was approximated as a continuous
covariate. The response variable ‘detection positive minutes per day’ (DPM) was
modelled with a quasipoisson distribution to account for over-dispersion. In addition,
an effort term was included in the model to account for any bias caused by difference
in the number of hours per day that the C-PODs were operational due to servicing.
Covariates ‘temperature (°C)’ and ‘productivity’ (chlorophyll-a, mg/m$^3$) were
included in the monthly model, and the response variable was modelled as detection
positive days (DPD) with a quasipoisson distribution again due to overdispersion of
the model residuals. Multi-collinearity of the covariates was tested for all of the models
by calculating Generalized Variation Inflation Factors, GVIFs (Fox & Weisberg
‘productivity’ and ‘temperature’ were modelled with cyclic cubic splines, and the
circular covariate, ‘Julian day’, with a b-spline. The different models and covariates
have been summarized in Appendix 5.1.

An autocorrelation function (ACF) plot was used to visually check the level of
temporal autocorrelation in the, and generalised estimating equations (GEEs; Liang &
Zeger 1986) were subsequently applied after fitting the GAMs with the purpose of
explicitly modelling the observed autocorrelation within the blocks (see also Dormann
et al. 2007; Pirotta et al. 2011; Pirotta et al. 2014; Culloch et al. 2016). GEEs can be
used to uncover marginal (i.e. population-averaged) effects rather than conditional
effects that relate to the estimated effect on an individual (Heagerty & Zeger 2000).
With the GEE approach, data points were divided into independent blocks, and a
correlation structure for the residuals specified within blocks (Liang & Zeger 1986).
One specific benefit of using GEEs in analyzing autocorrelated datasets is that they
allow for differences in the level of autocorrelation among blocks (Koper & Manseau
2009). In addition, parameter estimates and empirical standard errors in GEEs are
robust even in a situation where the correlation structure might be misspecified, as
found by simulation studies (Liang & Zeger 1986; Overall & Tonidandel 2004). Quasi-
likelihood under the model independence criterion (QIC) values (Pan 2001) were used to select the best autocorrelation structure between “working independence”, “exchangeable” and “AR1” structures in the binomial hourly model and in the daily model with Poisson distributed data. Repeated Wald’s tests were used to assess the significance of the retained covariates (Hardin & Hilbe 2003). Finally, diagnostic residual plots were inspected to assess the fit and predictive power of the best model.

5.3. Results

In total, out of more than 24,000 hours of deployment, the data set collected with both of the C-PODs included 6,600 detection positive minutes where at least one bottlenose dolphin echolocation click train was logged (Table 5.1). The number of detection positive days (DPD) per month in 2013–2015 are shown for both C-PODs in Fig. 5.2. The results show that dolphins were detected in Killary Fjord every month during the deployment with presence on at least 15 days during the months of March–July 2015 (Fig. 5.2), whereas dolphin presence was in general lower in McSwyne’s Bay with no detections logged during January–February 2015 (Fig. 5.2).

<table>
<thead>
<tr>
<th>Year</th>
<th>No. of deployment days</th>
<th>Detection positive days</th>
<th>Detection positive minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>McSwyne’s Bay</td>
<td>Killary Fjord</td>
<td>McSwyne’s Bay</td>
</tr>
<tr>
<td>2013</td>
<td>74</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>2014</td>
<td>314</td>
<td>108</td>
<td>81</td>
</tr>
<tr>
<td>2015</td>
<td>260</td>
<td>265</td>
<td>46</td>
</tr>
<tr>
<td>Total</td>
<td>648</td>
<td>373</td>
<td>141</td>
</tr>
</tbody>
</table>

Table 5.1 C-POD deployment summary showing the number of deployment days per year for each study site, and the number of detection positive days and minutes logged during those days.
**Figure 5.2** Detection positive days (DPD) per month logged by C-PODs during the deployment in 2013-2015. Note that the break in the line means that the C-POD had run out of batteries and was not logging clicks.

**Hourly model**

The average tidal current speed was 0.04m/s (SD = 0.02) in Killary Fjord, and 0.04m/s (SD = 0.05) in McSwyne’s Bay. The mode of tidal current direction was 290° in Killary Fjord and 292° in McSwyne’s Bay. Significant positive temporal autocorrelation was found in the hourly data set (Run’s test statistic: -66.79, $P < 0.001$), and after inspecting the correlation plot (Fig. 5.3), the data were divided into weekly blocks for which the autocorrelation was modelled using GEEs (Liang & Zeger 1986). The GEE-GAM with “working independence” autocorrelation structure was chosen over the “exchangeable” and “AR1” based on the QIC-values (“independence” QIC: 6993.6, “exchangeable” QIC: 7034.8, “AR1” QIC: 7011.1).
Significant explanatory variables kept in the best performing GEE-GAM were ‘daylight’ with an increase in the probability of detections during hours of daylight ($P < 0.05$), ‘water level‘ with the probability of dolphin detections increasing at mid and anf higher tidal levels ($P < 0.0001$), tidal current speed’ ($P < 0.0001$) with an observed bimodal effect on the probability of dolphin presence, and ‘tidal current direction’ ($P < 0.001$) with more detections occurring with northerly and southerly tidal flows (see Fig. 5.4 for the partial residual plots of these variables). The model fit and predictive power were poor with concordance correlation and marginal $R^2$ values between fitted and observed values of 0.020 and 0.011, respectively, implying that only very a small proportion of variation in dolphin presence was explained by these covariates.
Figure 5.4 Partial residual plots of the best GAM-GEE model for dolphin presence/absence and significant explanatory continuous covariates in the hourly model. (A) Estimated relationship with ‘water level’. (B) Estimated relationship with ‘tidal current speed’. (C) Estimated relationship with ‘tidal current direction’. The dotted lines represent the 95% confidence intervals, and the rug plot on the x-axis shows the actual data values, with probability of presence and absence on the upper and lower part of the plot, respectively.

Daily model

Significant positive autocorrelation (Run’s test statistic: -16.69, P <0.001) in the detection positive minutes was modelled by dividing the data into weekly blocks, for which the autocorrelation was modelled using GEEs (Liang & Zeger 1986) with a “working independence” correlation structure, based on the inspection of the GAM residuals (Fig. 5.5) and the QIC values (“independence” QIC: -13128.4, “exchangeable” QIC: -13127.2, “AR1” QIC: -13121.8). Significant covariates kept in the best model were ‘site’ with significantly more detections logged in Killary Fjord (P <0.001) and the interaction between ‘site’ and ‘Julian day’ with a bimodal response in McSwyne’s Bay (P <0.001, 5.6a) with a slight increase in detections at ~100 days (March) and a higher peak at ~270 days (September), and a single peak in the number of detections in Killary Fjord at ~120–150 days, i.e. in April–May (P
<0.001, Fig. 5.6b), and. The concordance correlation and marginal $R^2$ values between fitted and observed values were 0.080 and 0.043, respectively, indicative of poor model fit and predictive power.

Figure 5.5 Correlation of the daily GAM residuals for each block (grey lines) and mean autocorrelation (red line) at each lag. Autocorrelation values of >0.05 denote significant autocorrelation. The time lag is given in hours, 150h corresponds to approximately 6 days, after which autocorrelation diminishes.
Figure 5.6 Partial residual plots of the best GAM-GEE model for the detection positive minutes and the significant explanatory continuous covariate, ‘Julian day’, in (A) McSwyne’s Bay, Co. Donegal, and (B) Killary Fjord, Co. Galway, in the daily model. The dotted lines represent the 95% confidence intervals, and the rug plot on the x-axis shows the actual data values of counts.
Monthly model

The temperature varied from 6.9°C recorded in February of 2015 to 16.0°C in July of 2014, and productivity measured as chlorophyll-a concentration varied from 0.2 to 17.2 mg/m$^3$ in the area surrounding Killary Fjord, and from 0.8 to 37.7 mg/m$^3$ around McSwyne’s Bay. No temporal autocorrelation was found in the monthly dataset, therefore, GEEs were not incorporated in the GAM. From the covariates, only ‘mean productivity’ was kept in the best model but it was non-significant ($P = 0.190$).

5.4. Discussion

In addition to obtaining information on site occupancy and habitat use of bottlenose dolphins, understanding which environmental factors influence habitat use is important in the initial selection and subsequent monitoring of protected areas. In an attempt to answer these questions, echolocation click detections of bottlenose dolphins at two different locations on the west coast of Ireland were modelled with a number of environmental parameters that have the potential to influence the habitat use of dolphins on different temporal scales.

A significant increase in dolphin detections was found during daylight hours, which can either imply that dolphins were more likely to visit the sites during the day, or that they were more active in producing echolocation clicks during daylight hours. In contrast, a significant increase in occurrence and foraging activity (presence of feeding buzzes, click trains that have short and progressively decreasing inter-click intervals) was recorded during hours of darkness in a study with resident Ocean humpback (Sousa plumbea) and Indo-Pacific bottlenose dolphins (Tursiops aduncus) in Zanzibar, Tanzania (Temple et al. 2016). However, data collected in Sarasota Bay, Florida, have shown that bottlenose dolphins are acoustically active at night as well as during the day and that foraging activities are undertaken during both day and night with echolocation clicks and buzzes occurring throughout the entire 24-hour period (Wells et al. 2013). If the (acoustic) activity budgets of bottlenose dolphins on the west coast of Ireland are similar to the ones in Sarasota Bay, the scenario where the dolphins visit these coastal sites more during daylight hours rather than a difference in echolocation production seems more likely, assuming that they are producing echolocation clicks when they are within the effective detection range (300-400m, Nuuttila et al. 2013) of
C-PODs. Even though echolocation has been found to occur significantly less frequently when bottlenose dolphins are travelling compared to other activities (Jones & Sayigh 2002; Dos Santos & Almada 2004), echolocation click detections have been found to corresponded well with visual detections, with over 80% of the dolphin schools observed within 500m also detected acoustically (Philpott et al. 2007) indicating a high probability of detection due to active echolocation behaviour.

Tidal height and current speed had a significant effect on the number of detection positive minutes with a peak in detections at mid to high water levels and intermediate and fast current speeds. Similarly, in another study, bottlenose dolphins were found to be significantly more abundant during flood tide (i.e. incoming tide), particularly during the stationary stage of the tidal front (Mendes et al. 2002). Pirotta et al. (2014), however, found no significant effect of current speed on the presence of feeding buzzes. However, occurrence of feeding buzzes was not the focus of this study, and Pirotta et al. (2014) did not investigate how well the presence of feeding buzzes corresponded to the presence of echolocation clicks; thus the results in these two studies are not directly comparable. Nevertheless, the significant effect of tidal height and current speed found in this study may be influenced by different tidal flow conditions between the study sites, which in turn may facilitate the presence of different types of prey. It may be that the increased presence of dolphins associated with higher water level and faster current speed found in this study reflects the movements of prey species (Sveegard et al. 2012; see review by Benjamins et al. 2015). Telemetry studies, although concentrating on larger species such as blue marlin, *Tetrapturus audax*, or blue shark, *Prionose glauca*, have found an effect of currents on fish swimming speed and directionality (Carey & Scharold 1990; Brill et al. 1993).

Interestingly, Atlantic salmon (*Salmo salar*) smolts have also been shown to take advantage of tidal ebbs during their seaward migration and hold their position during flood tides (Moore et al. 1995; Lacroix and Curdy 1996). In addition, adult salmon have also been shown to time their movements with tidal currents in estuaries when returning to breed (Potter 1988). Atlantic salmon is likely to be an important part of bottlenose dolphin diet in Ireland, at least seasonally. Adult fish have been recorded from stomach contents of dolphins stranded around Ireland (Hernandez-Milian et al. 2015) and shown to be an important resource for a nearby resident population of bottlenose dolphins inhabiting Moray Firth, Scotland (Janik 2000; Bailey & Thompson 2010).
Further, anecdotal evidence exists that salmon fishery catches are higher during spring tides when the current speed reaches its peak. Concordant to the findings by Pirotta et al. (2014), tidal current direction was also significant in this study with an increase in echolocation bouts being recorded at currents running from South to North or from North to South (Fig. 5.4b). The tidal flood on the west coast of Ireland has a northerly direction (Anonymous 2004), and this could also influence movements of Atlantic salmon or other prey species.

In addition to tidal parameters, an increase in dolphin echolocation click detections in the spring was also recorded in this study in Killary Fjord. Pirotta et al. (2014) found an increase in feeding buzzes coinciding with the summer months in Moray Firth, Scotland. Even though the present study did not distinguish and classify feeding buzzes within the echolocation click bouts, it is likely that they are represented widely in the dataset. In fact, evidence exists that echolocation is used more frequently for feeding than in any other behavioural context (Jones and Sayigh 2002; Nowacek 2005; Gannon et al. 2005). Unlike the more northerly parts of Europe, where the timing of the runs of Atlantic salmon is largely dictated by temperature and melting of ice and thus limited to June–September, salmon can be found migrating back up river to spawn almost on any day of the year around the British Isles and Ireland (Sutterby & Greenhalgh 2005, Reed et al. 2016). Run timing is one of many life-history traits used to characterise Atlantic salmon, and differences in salmon migration between rivers and even between tributaries has been known for a long time. This variation in run timing is typically associated with time spent feeding at sea with larger multi-sea-winter individuals, which spend two or more years at sea, tend to enter rivers earlier in the year (spring) than smaller fish that have spend only one year at sea (Reed et al. 2016). However, the run peaks predominately occur during the spring (March–April) and summer (June–July) in Irish rivers (Quinn et al. 2006). It is thus possible that these peaks in salmon runs, especially the spring peak which is dominated by larger multi-sea-winter individuals (Quinn et al. 2006), are the drivers for dolphin presence in certain coastal sites and explain the greater number of detections coinciding with spring months in the Killary Fjord. Increased presence of dolphins linked to salmon run times has also been suggested by other authors (Wilson et al. 1997; Mendes et al. 2002), and is supported by the fact that River Bundorragha, a tributary of Killary Fjord, contributes to substantial number of spring salmon catches (Reed et al. 2016).
Significantly more detections were logged by the C-POD deployed at the mouth of Killary Fjord, Co. Galway than by the C-POD in McSwyne’s Bay, Co. Donegal. Killary Fjord and the associated tributaries are some of the most important salmon rivers in Ireland, and this combined with a large number of visual encounters indicates that the fjord is an important foraging area for bottlenose dolphins. Moreover, this is further supported by numerous observations of bottlenose dolphins chasing and capturing Atlantic salmon within the fjord. It may be that hydrographic features such as steep and narrow channels in the bottom of the fjord combined with currents resulting from tidal flow may gather and concentrate prey (salmon and other) thus facilitating capture. The deployment site in McSwyne’s Bay, Donegal, on the other hand, is a much wider and more open bay with different bathymetry and hydrographical features compared to Killary Fjord. It is located ~25km north from the closest salmon river, River Eske, and this may be one reason explaining the fewer number of detections logged at this site and the bimodal peak in detections occurring during the winter months.

Temperature was not found to be a significant factor explaining the presence of echolocation click trains logged by the C-PODs, and primary productivity (approximated as chlorophyll-a concentration, mg/m³) was only marginally significant with detection positive days decreasing with the amount of primary productivity (Fig. 6). Conversely, Hartel et al. (2015) found bottlenose dolphins in northern New Zealand to utilize deeper waters during the summer months and shallower waters in the winter months, and suggested that temperature associated with prey availability could be a possible factor explaining the difference in the fine-scale habitat use.

Even though the false positive rate reported with C-PODs is generally very low, between 1–4% (Nuuttila et al. 2013; Roberts & Read 2015), it is possible that some logged clicks were produced by other dolphin species than bottlenose dolphins. However, the likelihood of occurrence of this kind of false positives can be minimised by placing the devices in locations that are rarely, if ever, used by other species than the target species. Other dolphin species such as the common dolphin, *Delphinus delphis*, a relatively pelagic species (Perrin 2002; Bearzi et al. 2003), have never been observed in Killary Fjord, and their occurrence in McSwyne’s Bay is rare; thus it is likely that the false positive rate due to other species is very low in this study. It has been shown that C-PODs are efficient at detecting echolocation bouts within a radius
of almost up to 1800m (Nuu tila et al. 2013). However, the GENENC click train classifier used in this study has, in general, a tendency for false negatives rather than false positives (Roberts and Read 2014; Robbins et al. 2015). This coupled with the fact that C-PODs use click trains to identify the occurrence of echolocation and thus will not detect events containing less than five successive clicks with similar inter-click intervals (Tregenza 2013), and that echolocation clicks are highly directional and off-axis click trains might be missed, may lead to an underestimation of dolphin occurrence. Nevertheless, due to their low false positive rate and efficacy in detecting click bouts, C-PODs can be considered as an efficient monitoring tool in predicting dolphin presence particularly in sites where habitat use is sporadic and unpredictable. High mobility coupled with largely unknown ranging behaviour which can be driven by a suite of environmental or biological factors can present unprecedented challenges to the efficient monitoring of populations. This study demonstrates the potential of using passive acoustic monitoring devices, such as C-PODs, to unveil useful information on seasonal and temporal habitat use and can be used in conjunction with environmental factors to examine variables affecting the use of specific sites by bottlenose dolphins. The study also suggests that C-PODs can be used as a long-term monitoring tool in a relatively cost effective way that could potentially cover part of the management and monitoring requirements set by the EU Habitats Directive and the Marine Strategy Framework Directive.
References


activities and vessel traffic on marine mammals. Marine Ecology Progress Series 549, 231–242. doi:10.3354/meps11686


Chapter 5: Appendices

Appendix 5.1 Details of GAMs or GEE-GAMs ran on the C-POD data

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hourly</td>
</tr>
<tr>
<td>Response variable</td>
<td>Presence/absence of click trains</td>
</tr>
<tr>
<td>Distribution of response</td>
<td>Binomial</td>
</tr>
<tr>
<td>Covariates tested</td>
<td>Daylight (f)<em>, tidal level</em>**, current speed***, current direction**</td>
</tr>
<tr>
<td>Autocorrelation?</td>
<td>Yes → GEEs</td>
</tr>
<tr>
<td>No. observations</td>
<td>24449</td>
</tr>
<tr>
<td>No. independent blocks</td>
<td>154</td>
</tr>
<tr>
<td>Model predictive power</td>
<td>2%</td>
</tr>
</tbody>
</table>

(f) denotes a factor variable
Significance levels *** <0.0001, ** <0.001, * <0.05
Chapter 6: Concluding discussion

Commonly used methods in conservation and management in the context of the present study

Protecting species and their habitats is the goal of conservation biology, and this could not be achieved without efficient management strategies. From the several methods currently applied to the assessment and management of cetacean populations, The International Union for the Conservation of Nature (IUCN) Red List Criteria classifies species to different categories on the basis of their abundance in relation to their risk of extinction: in short, a species with fewer than 50 mature individuals is classified as ‘Critically Endangered’, one with fewer than 250 classified as ‘Endangered’, and fewer than 1000 classified as ‘Vulnerable’ (IUCN 2016). In addition, species can be classified as of ‘Least Concern’ or ‘Near Threatened’ if they are not considered to be under immediate threat, and if sufficient data on their abundance exists (IUCN 2001). However, even though IUCN uses panels of experts to weigh the status of individual species against a Population Viability Analysis, it does not make any direct management recommendations (Lonergan 2011).

The European Union's Habitats Directive (Council Directive 92/43/EEC), on the other hand, requires each of the EU Member States to maintain or restore all the marine mammal species in European waters, at a ‘favourable conservation status’. This status is reached when population dynamics data of a species indicate that populations are maintained at a viable level in the long-term in their ‘natural habitat’, the species’ natural range is neither being reduced nor is likely to be reduced for the conceivable future, and a sufficient amount of habitat of suitable quality exists to maintain its populations on a long-term basis (Council Directive 92/43/EEC). Each Member State is required to report on the status within its boundaries every six years, however, it is up to the country to decide on the way it adheres to the requirements (as this is a Directive and not a Regulation). This approach has been criticised on the basis that the results of these reports are likely to be biased due to the lack of a common set of monitoring methods used in every EU Member State (Lonergan 2011). The Directive has further guidelines to defining the status of the species/populations and uses
‘Favourable Reference Values’ as baseline (Habitats Directive/Article 17); species that had a “normal” age structure and were above their “favourable reference population” were classed as ‘Favourable’; those with a structure that “strongly deviated from normal”, or were 25% below the “favourable reference population”, or were below it and had declined by 1% annually in the previous six years were to be classified as ‘Unfavourable – Bad’, and all others ‘Unfavourable – Inadequate’. However, one criticism raised for these guidelines has been that they do not consider the natural variation in population sizes when setting the initial ‘Favourable Reference Values’ (Lonergan 2011).

The EU Marine Strategy Framework Directive (MSFD, Council Directive 2008/56/EC) has the general goal to “achieve ‘Good Environmental Status (GES)’ of EU marine waters by 2020 and to protect the resource base upon which marine-related economic and social activities depend” (www.ec.europa.eu), thus it follows an ecosystem-based approach in the management of human activities whilst promoting sustainable use and protection of the environment. According to the MSFD, each EU Member State is obliged to develop an ‘adaptive management strategy’ for its marine waters, and this strategy needs to be reviewed every six years thus making the management protocols more dynamic and, at least, hopefully more responsive to any changes that may occur in population dynamics of protected species. In addition, the Member States are encouraged to co-operate regionally when developing marine strategies; the co-operation of different regions is coordinated through Regional Sea Conventions, for example in the case of Irish and UK waters, the relevant regional convention is the Oslo-Paris (OSPAR) Convention. The monitoring strategy within the MSFD developed by Ireland and involving the indicator species for GES (including bottlenose dolphins), is currently under review by OSPAR, but will be based on monitoring under the Habitats Directive (MSFD/Article 11). The MSFD’s integrated approach towards ecosystem-based management is a welcomed idea along with its encouragement for co-operation between the Member States thus promoting the notion of network of SACs within EU waters. However, this approach will have to be backed up by clear and comprehensive management protocols. It remains to be seen how integrative these protocols, that are currently under development, will be.

Whereas the Habitats Directive (and possibly also MSFD) falls short in the sense that it does not provide specific guidelines to the Member States in the monitoring of the
status of species listed in Annex II and IV, in other jurisdictions, such as the USA, legislation such as the Marine Mammal Protection Act (1972) (MMPA) sets more rigorous and quantitative goals with the overall aim of preventing the depletion of local populations of marine mammals (referred to as stocks), and to restore them to level of Maximum Sustainable Yield (MSY). Later on, ‘Potential Biological Removal’ (PBR) was added to the Act (Wade 1998) as a way to quantitatively assess the effect of anthropogenic impacts on populations. The MMPA gives taxon/species specific guidelines regarding stock assessment (Wade et al. 1997), but in general, the purpose of the assessment is to determine the level of mortality that the stock can sustain. The calculations usually require several types of information, such as current and historical abundance, estimates of age of maturity, spatial distribution, rate of natural mortality, pregnancy rate (i.e. inter-birth interval), age distribution and MSY (Breiwick & York 2009). In addition to more comprehensive monitoring guidelines of the MMPA, the status of coastal bottlenose dolphin stocks is reported every 1–3 years in the US (http://www.nmfs.noaa.gov/pr/sars/species.htm); this is twice as often as required by the EU’s Habitats Directive.

The main goals of this dissertation were:

1) to delineate the populations occurring in Irish waters as Management Units (MUs), to model the colonisation history of bottlenose dolphins in the wider North Atlantic Ocean with an emphasis on the colonisation of the coastal northern latitude habitats, and to describe and discuss some of the driving forces shaping the observed population structuring (Chapters 2 and 3)

2) to derive an abundance estimate for the wide-ranging ‘coastal mobile’ population and to estimate the scale of movements of this population (Chapter 4)

3) to describe the spatio-temporal variation in dolphin habitat use in the context of a number of environmental factors (Chapter 5), and

4) to discuss the findings in relation to the management of populations (Chapters 3, 4 and 5).

In Chapter 2 on phylogenetics and biogeography, the results indicated that the coastal bottlenose dolphins currently inhabiting the northern parts of the NE Atlantic may have
originated from refugial population(s) occupying the Mediterranean Sea during the Last Glacial Maximum via founder event(s), and that the colonisation of the NE Atlantic likely occurred following deglaciation after the LGM (see Fig 6.1.). This phylogenetic analysis includes, to the best of my knowledge, the first attempt to overcome the effects of time-dependency on the nucleotide substitution rate through a combination of using only third codon positions (sites where most mutations are silent, or synonymous, and are likely to be retained) of the coding genes of the mitochondrial genome and incorporating both tip calibration points from complete dated ancient mitogenomes and a fossil calibration on a deeper node. The estimates for nucleotide substitution rates, coalescence times and clade crown ages obtained in this study are in agreement with previously published estimates (e.g. Duchene et al. 2011; Moura et al. 2013; Morin et al. 2015), as well as the climatological and geological time frame, for example, the opening of the Bosphorous Strait and the retreat of the ice sheets covering Northern Europe.

Figure 6.1 The most likely colonization patterns of bottlenose dolphins to the coastal NE Atlantic estimated in this study. (A) Clade consisting of samples from coastal West Ireland, West Scotland and Brittany. (B) Clade consisting of samples from coastal West Ireland, East Scotland, England, Wales and Brittany.

Climatic oscillations are thought to have played a role in shaping species distribution and divergence (Avise & Walker 1998), and combining population genetic data with models for suitable habitat such as the AquaMaps may offer a way to predict responses of current populations to the ongoing climate change, a field of research where only a
few examples exist (e.g. Inoue & Berg 2016). Data from previous genetic studies on marine mammals (Pastene et al. 2007; DeBruyn et al. 2009; Amaral et al. 2012; Foote et al. 2013) and other marine species (Crandall et al. 2012) have suggested that changes in population structuring and connectivity have followed glacial cycles. From this study, it also appears that cladogenesis events were correlated with periods of temperature change, with warmer temperatures at the onset of Holocene leading to an increase in sea-level in coastal areas and the subsequent release of available habitat coinciding with rapid radiation and population expansion of bottlenose dolphins into northern latitudes. These kind of rapid ‘leading edge expansions’ usually resulted in reduced genetic diversity in temperate species in large areas in the northern parts of Europe (Hewitt 1999) with a number of studies showing greater homozygosity in northern expansion areas (Hewitt 1996), and the northernmost coastal populations inhabiting the waters around Scotland and the west of Ireland are likely to be examples of this leading edge. In contrast, slower expansion and varied topography in southern European latitudes would, in general, retain more genetic diversity in southern populations allowing more time for divergence over many glacial periods in various southern refugia whereas northern temperate populations would die off during these colder periods (Hewitt 1999). This phenomenon is also documented in the current coastal bottlenose populations of the NE Atlantic, with the ‘Coastal South’ population having more genetic diversity in nuclear (Louis et al. 2013a) and mitochondrial markers (see Chapter 2).

The lack of differentiation of the ‘Shannon’ and ‘mobile’ bottlenose dolphin samples into separate haplogroups, or clades (Fig. 2.4, Chapter 2), indicates a recent population divergence that was found based on nuclear markers (Mirimin et al. 2011; Louis et al. 2014a, Chapter 3). In fact, Louis et al. (2014b) estimated the timing of divergence between ‘Coastal South’ and ‘Coastal North’ populations to have happened ~2560 yBP, and it is possible that the bottlenose dolphins resident in the Shannon estuary diverged from the ‘mobile’ population even more recently.
The work carried out for this PhD project has contributed to the assessment of the status of bottlenose dolphins in Irish waters. Specifically, in Chapter 3, further evidence was provided for the existence of three distinctive populations in Irish waters in support of Mirimin et al. (2011). This has been shown by analysing a larger set of genetic samples collected on a wider coastal area, and also by applying a range of statistical methods, including kinship-based methods and randomisation tests (e.g. Palsbøll et al. 2010). In addition, the photo-identification data in this study provided the first evidence for the near complete social isolation\(^1\) of the two coastal populations, previously documented to exist only between the coastal and putative pelagic populations (Oudejans et al. 2015). For the first time, the recent (over the past two generations) genetic dispersal between the populations occurring in Irish waters was quantified, and it was established that the three populations are effectively genetically isolated and thus should be defined as separate Management Units (MUs) (see Chapter 1; Moritz 1994; Wade et al. 1997; Waples & Gaggiotti 2006; Palsbøll et al. 2006).

Geographic isolation in new environments may contribute largely to divergence (Wright 1942). This is not likely to apply to the two populations of coastal bottlenose dolphins in Irish waters given their adjacent ranges, so the underlying reason behind the apparent reproductive isolation is likely to be driven by a combination of socio-ecological factors. More research effort is needed to fully elucidate the drivers of this existing population structure, but the evidence from this study infers that at least site-fidelity and social associations may be contributing factors. These reasons have been also suggested by several other authors to drive population structuring among several marine and terrestrial species (e.g. Parsons et al. 2006; Lowther et al. 2012; Louis 2014a,b; Podgórski et al. 2014; Gaspari et al. 2015). Long-term site-fidelity lasting over several years coupled with long and short term social bonds and higher relatedness (compared to the offshore population, although this may be an artefact of smaller proportion of the offshore population being sampled) was found within the two coastal populations. Prey resources are temporally or spatially more predictable in coastal estuarine environments compared to open ocean habitats, and in these areas different foraging strategies may be socially and culturally transmitted (e.g. Mann & Sargeant 2003; Krützen et al. 2005) potentially leading to genetic divergence (see

\(^1\) An individual sighted in McSwyne’s Bay, Co. Donegal, in 2014 was photographed in the Shannon estuary in the summer of 2015. This dolphin was recorded associating with members of the Shannon community. However, the Shannon photo-id data from 2015 was not analysed for this thesis.
Kopps et al. 2014). However, some areas requiring more research effort were identified during the course of this study; for example, more tissue samples from ‘pelagic’ animals are needed in order to investigate whether further structuring exists within this population. Increased sampling of pelagic animals would also help to provide a more rigorous understanding of the distribution and habitat use of this population in contrast to the two coastal populations. In addition, in order to effectively assess sex-biased dispersal, more samples from females need to be collected from coastal populations, as biopsy sampling in Irish waters so far has been concentrated largely on males possibly due to their distinctive and more obvious markings. Nevertheless, from a management point of view, the fact that these coastal bottlenose dolphins showed non-overlapping ranges with site-fidelity to each of the coastal SACs, implies that the designation of these areas to correspond with the different populations has been successful.

Following the delineation of MUs, the abundance of the ‘coastal mobile’ population was estimated in Chapter 4 using a relatively novel Bayesian multi-site method in mark-recapture abundance estimation. Bayesian methods have especially been applied to abundance estimation of cetacean species (e.g. Durban et al. 2005; Durban et al. 2010; Moore & Barlow 2011; Fearnbach et al. 2012; Cheney et al. 2013) as they are well-suited for sparse data sets. The Bayesian multi-site approach is particularly useful when it is not feasible to apply a survey strategy covering the entire distribution of the population in question (Durban et al. 2005). This method can be used relatively economically with the help of a sightings network when the presence of the animals is unpredictable and spatially patchy and the surveys are largely limited by budget and weather conditions. By using this multi-site approach, where re-captures of individuals were ordered on a spatial rather than temporal scale, an abundance estimate was derived that is both accurate and robust for the wide ranging ‘coastal mobile’ population of bottlenose dolphins that use the West Connacht Coast SAC. Deriving an estimate using only data collected within the designated area would have likely resulted in an underestimation of abundance due to the potential failure to capture some of these highly mobile animals. In addition to abundance estimation, it was determined that six consecutive years of annual monitoring would be required to detect a 10% annual decline in population, even with a low CV of 0.11. Therefore, it is recommended that monitoring should be continued on a yearly basis. Moreover, in
order to be able to detect an overall decline of 25% in abundance over a six-year reporting period, a limit set by the Habitats Directive, the CV around the abundance estimate would have to be as low as 0.08, which could be difficult, if not impossible, to achieve for a mobile population such as the one in this study. As a conclusion, continued and more wide-scale research effort is recommended in order to detect any changes in population dynamics within the coastal populations, and to better understand their ranges. Shorter reporting interval of 1–3 years used in the management of the coastal bottlenose dolphin ‘stocks’ in the US, combined with a thorough assessment of mortality and fecundity rates, should be also applied to the management of coastal populations within European waters.

The Bayesian multi-site method offers great potential to be used as a monitoring tool for networks of MPAs, such as the national SACs that form part of a European-wide conservation instrument directed to protect populations on a wider transboundary level. More international co-operation, encouraged by the MSFD and to some extent also the Habitats Directive, is required in the assessment of the status of the coastal mobile bottlenose dolphin population, especially given that evidence suggests that these mobile animals have a range extending beyond country boundaries (Robinson et al. 2012), and that they may belong to a wider meta-population of dolphins occupying the waters of Ireland, Scotland and northern France (Nichols et al. 2007; Louis et al. 2014a). In fact, plans for a multi-national project to assess the population status of bottlenose dolphins in countries of the “Atlantic Arc” (Ireland, UK, France, Portugal and Spain) are under way. In the meantime, a smaller scale project resolving the (meta)population status of the bottlenose dolphins in Irish waters could be done by analysing a common set of microsatellite markers used by Louis et al. (2014a) or the SNP markers generated in a recent study (see Louis, Nykänen et al. in prep, Appendix 6.1).

Studying how individuals use areas on a smaller spatial scale is a question that also poses logistical difficulties and is often confined to summer months in northern latitudes due to weather conditions alone. For this thesis, site occupancy of bottlenose dolphins within a location inside the West Connacht Coast SAC was examined using passive acoustic monitoring (Chapter 5), and compared to site occupancy in another location over 100km north outside the SAC boundaries. The effect of environmental factors that are likely to influence habitat use of bottlenose dolphins were also
examined. Killary Fjord, Co. Galway, was identified as an important site for the dolphins with monthly detections and significantly more detection positive days logged in this site compared to McSwyne’s Bay, Co. Donegal. This indicates that the area designated as a bottlenose dolphin SAC containing the area around Killary Fjord is indeed a site regularly used by dolphins. A peak in detections occurred in the spring in this site which may be a response to a local increase in prey availability. These results show that assessing the effectiveness of SACs can be supplemented by PAM, and it is recommended that monitoring be continued in order to gather long-term evidence of the effect of seasonal trends in relation to dolphin occupancy, as only 13 months of continuous data from both sites during this study were collected. In addition, more C-PODs should ideally be deployed in areas within the SAC as well as outside, in order to identify other areas of importance and seasonal factors effecting site occupancy. Moreover, having a strategically placed network of C-PODs along the coast of Ireland could increase the likelihood of detecting changes in the site occupancy and help to re-evaluate the status of the current SAC.

**Areas for future research**

In general, this thesis demonstrates that applying a combination of methodological approaches and selecting these approaches to suit the populations in question can provide an efficient way to monitor mobile populations with unpredictable ranges and habitat use such as the transient population of bottlenose dolphins on the west coast of Ireland. However, some areas in which further research is required were identified are discussed below.

Understanding the spatial distribution of organisms needing protection is one of the key aspects for effective management and conservation (Whittaker *et al.* 2005). The full ranging patterns of these highly mobile animals (i.e. the ‘coastal mobile’ population) is not known; however, evidence exists that at least some identified individuals range widely and beyond trans-national boundaries. One solution to overcome this challenge caused by high mobility and unpredictable movement patterns could be to increase the collaboration between researchers in different countries and areas. This would include sharing of the photo-id catalogues and tissue samples between research institutes. Repeated wide-scale survey effort is essential in
order to be able to detect shifts in distribution of these animals and also to detect changes in their abundance. Another approach to investigating animals’ ranging behaviour would be to conduct a telemetry study using remotely implanted tags attached to individual dolphins. Satellite tagging has the advantage of providing continuous movement and range information over the period the tag is attached and transmits data. A disadvantage is that it would only be possible to tag a few individuals due to the cost and the logistical difficulties involved in the tagging. However, this approach would be highly complementary when combined with photo-id and genetic mark-recapture methods. At present there are ethical considerations for such a tagging programme, but tags are becoming smaller and more readily employed using remote methods such as rifles and crossbows (e.g. LIMPET-tags, Wildlife Computers Inc.) instead of more conventional capture tagging methods.

More effort should also be invested into estimation of demographic parameters such as survival, mortality rate, the calving rate and the proportion of adults/subadults in the population. Most higher predators have long life spans, and consequently it can take several years before any changes in population growth or structure become apparent; again this emphasises the need for regular monitoring. A potentially promising method in detecting changes in fecundity and survival within a population would be to look at changes in age structure (Holmes & York 2003). Recently, it was found that fluctuations in reproductive rates can have considerable impacts on population viability of bottlenose dolphins (Manlik et al. 2016), thus again highlighting the importance of consistent monitoring of these parameters. Consequently, models of population viability could be run including different population parameters, such as the ones listed above and following Manlik et al. (2016), and with different levels of prey resources (e.g. depleted fish stocks).

Other future research could also include comparison of genomic differentiation and gene expression between the two distinct coastal populations, to identify functional regions, and to subsequently investigate the role of different genes in local adaptation (see review by Kelley et al. 2016). In addition, greater geographic resolution in resolving NE Atlantic bottlenose dolphin phylogeny could be obtained by performing a larger scale analysis with a wider coverage of samples from the pelagic and other coastal northern European populations. In fact, twenty more modern samples collected from the coastal areas of Scotland and three radio-carbon dated ancient subfossil
samples originating in the Dutch Southern Bight in the North Sea will be sequenced in
the near future and analysed with the dataset presented in Chapter 2.

A comprehensive estimate of diet combined with genetic sampling is required to
investigate the role of resource partitioning as a driver of species or population
structure (e.g. Foote et al. 2009; Kiszka et al. 2011; Ryan et al. 2013; Ansmann et al.
2014; Louis et al. 2014b). Hernandez-Milian et al. (2015) examined the stomachs of
12 bottlenose dolphins stranded around the Irish coast, but these samples have not all
been genotyped. The investigation of stable isotopic signatures has the potential to
reveal differences related to population structure on a wider scale, for example, the
coastal – offshore separation (e.g. Louis et al. 2014b; Rogan et al. in prep.) or
sympatric species or populations with overlapping distributions (e.g. Foote et al. 2009;

**Utility of Marine Protected Areas in the conservation of marine mammals**

According to the definition by the World Conservation Union (IUCN 1994), a Marine
Protected Area (MPA) is “any area of intertidal or subtidal terrain, together with its
overlying water and associated flora, fauna, historical and cultural features, which has
been reserved by law or other effective means to protect part or all of the enclosed
environment”. MPAs have recently evoked discussion on their usefulness in
preserving biodiversity or protecting a specific species or population (e.g. Hooker &
Gerber 2004; Agardy et al. 2011; Hartel et al. 2015; Wilson 2016). One criticism has
been the argument that the establishment of MPAs is sometimes solely driven by the
interest shown by the public to charismatic species like marine mammals (Hooker &
Gerber 2004; Wilson 2016), and they have sometimes been referred to as ‘paper parks’
that lack the necessary monitoring and regulation (Duffus & Dearden 1995; Hooker et
al. 1999; Agardy et al. 2011). One example of this would be the designation of so
called whale sanctuaries that extend to cover the entire EEZs of countries (Rogan &
Berrow 1995; Hoyt 2005; Agardy et al. 2011). On first thought this idea seems very
appealing with the general aim of promoting the protection of marine mammals,
however, the question remains whether it is appropriate to call an area truly protected
when these designations are not accompanied by risk assessment, mitigation and a set
of specific management actions.
Marine Protected Areas have also been criticised for having been designated in areas that hold less importance for the species under protection whilst leaving more important areas unprotected due to various political or economic reasons. For example, the Pelagos Sanctuary in the Mediterranean Sea covering areas in waters of Italy, France and Monaco provides protection to areas that are of relatively low value for marine mammals whilst leaving out more important areas due to potential difficulties in managing these areas (Agardy et al. 2011). When protection is directed to a specific population, the ideal situation would be to ensure protection over the entire population’s range which it inhabits year-round (Reeves 2000). However, some mobile marine predators, such as the sperm whale, have a global population structure possibly through male-mediated gene flow and a distribution that extends from feeding grounds in low latitudes to calving area in high latitudes (Lyrholm & Gyllensten 1998; Alexander et al. 2016). The designation of an MPA to cover such a large offshore area would not be economically or politically feasible in most parts of the world. Thus one of the criticisms towards MPAs has been that they are too small and represent only a minute portion of the total range of the species (Hooker & Gerber 2004; Agardy et al. 2011; Wilson 2016). As an example, the generally applied criteria used in the designation of MPAs for marine mammals has primarily been concentrated on preserving breeding areas without taking to account foraging habitats or migration routes (Hooker & Gerber 2004). Yet, it is likely that most marine mammals are at most risk while foraging (e.g. harbour porpoise, ASCOBANS 2012). However, Hooker and Gerber (2004) also argued for the general benefits of MPAs that are designated based around vulnerable life stages, such as breeding/calving areas, by stating that even if the area were used by a species only for a portion of its life span, this would still diminish the overall cumulative impact of other threats, thus reducing the frequency with which each individual was exposed (Hooker & Gerber 2004). The bottlenose dolphin SACs in Irish coastal waters have been designated based on core ranges estimated from encounters during photo-identification surveys and considering the two distinct populations. However, McSwyne’s Bay was identified as an important site for bottlenose dolphins during previous surveys (Ingram et al. 2001) and this study, suggesting that this area should be considered for further SAC designation.

A considerable amount of thought has to be put into the management of MPAs. For example, managers are often too concerned about managing an MPA alone without
paying attention to the management of the entire population. Concentrating all monitoring efforts on a designated area that covers only a part of a mobile and wide ranging species, could give a biased view of the status of the population, if its range has shifted to other areas (e.g. Wilson *et al*. 2004). An economic solution to overcome this problem could be the deployment of a network of passive acoustic monitors to monitor the site use within the protected area and the wider area around it. Other methods could include the use of novel multi-site methods in abundance estimation that incorporate estimates on the scale of movement of animals within and outside of a MPA. Both of these methods were used in this study.

There are, however, some positives in amongst all the criticism that MPAs have received. For example, while most existing protected areas are isolated and thus connectivity between sites is not ensured, the Great Barrier Reef Marine Park in Australia has been described as a success story of a large scale network of MPAs with its integrated and adaptive management (e.g. McCook *et al*. 2010). Even when a MPA is designated to cover a relatively small isolated area, it can be successful, as in the case of the Banks Peninsula Marine Mammal Sanctuary designated for Hector’s dolphins (*Cephalorhynchus hectorii*). A recent study reported a 5.4% increase in mean survival probability following the designation of the MPA accompanied with a ban on commercial and a restriction in amateur gillnetting in the area (Gormley *et al*. 2012).

The Natura 2000 network of SACs for bottlenose dolphins at the European scale seems to be a step to the right direction because a network of SACs may ensure the protection of areas important for the dolphins and enhance connectivity between the SACs. However, transnational co-operation in the monitoring of these areas is required since the Member States are responsible for reporting on the status of species only in their own national SACs, and populations can have ranges extending beyond country boundaries. Hopefully the OSPAR convention will provide a solution by producing a comprehensive and effective management protocol that the Member States will then adhere to.

Static MPA boundaries may not be the most appropriate method to manage marine mammals. More dynamic MPAs where the boundaries can be adjusted in response to changing species distributions or site use have been suggested by several authors (Hooker & Gerber 2004; Hooker *et al*. 2011; Hartel *et al*. 2015). How this will be accepted by managers is, however, uncertain since the practicalities involved in
shifting the boundaries and management of these dynamic areas may turn out to be logistically problematic. Another approach suggested by several authors has been the development of more comprehensive marine spatial planning and ecosystem based management strategies (e.g. McLeod et al. 2005; Halpern et al. 2009; Agardy et al. 2011; Wilson et al. 2016) that emphasise the protection of the ecosystem as a whole whilst acknowledging the connectivity among systems (McLeod et al. 2005). The European Union’s Marine Framework Directive is an example of this (with bottlenose dolphins being listed as one of the indicator species of good environmental status of coastal habitats), but it remains to be seen how the management will be applied to these habitats in order to ensure the protection of the populations.

In conclusion, it is positive and forward thinking that the different populations with their non-overlapping core ranges have been taken into consideration in the designation of the SACs in Irish coastal waters, even though they do not cover the entire (largely unknown) ranges of the populations. Based on the results of the photo-id and acoustic monitoring work in this study, it seems that the areas around Killary Fjord are important for the dolphins. However, the area in Donegal Bay outside the SAC boundaries also seems important based on the regular encounters of large groups consisting of up to 100 dolphins (see Chapter 4). More research effort will be required to uncover the distribution of these animals with the possibility of extending the existing SACs to cover areas further north. These areas should at least cover the core ranges of the dolphins. Efficient and regular monitoring of the populations should be continued so that any changes in population parameters can be detected and the best possible conservation strategies implemented in a timely manner, ensuring the long-term viability of the populations.

References


Discrete or not so discrete: Long distance movements by coastal bottlenose dolphins in UK and Irish waters. Journal of Cetacean Research and Management.


Wilson, B., 2016. Might marine protected areas for mobile megafauna suit their proponents more than the animals? Aquatic Conservation: Marine and Freshwater Ecosystems 26, 3–8. doi:10.1002/aqc.2619


Chapter 6: Appendices

Appendix 6.1. Manuscript of SNP discovery (in prep.)

High density, genome-wide SNP discovery in Northeast Atlantic bottlenose dolphins based on genotype likelihoods from multiplex shotgun sequencing data

MARIE LOUIS,1,2,3* MILAJA NYKÄNEN,4* SHYAM GOPALAKRISHNAN,5 ANDREW BROWNLOW,6 WILLY DABIN,1 ROBERT DEAVILLE,7 EILEEN DILLANE,4 FRANCOIS GALLY,8 SIMON N. INGRAM,9 VALENTINA ISLAS VILLANUEVA10, KIM MAGNUSSEN,11 EMER ROGAN,4 BENOIT SIMON-BOUHET2, NAGARJUN VIJAY,12 NATHAN WALES5 and ANDREW D. FOOTE13

1Observatoire PELAGIS, UMS 3462 CNRS/Université de La Rochelle, Pôle Analytique, 5 allées de l’Océan, 17000 La Rochelle, France, 2Centre d’Etudes Biologiques de Chizé, CNRS-UMR 7372, 79360 Villiers-en-Bois, France, 3Scottish Oceans Institute, University of St Andrews, St Andrews, Fife, KY16 8LB, Scotland, 4School of Biological, Earth and Environmental Sciences, University College Cork, Ireland, 5Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark, 6Scottish Marine Animal Stranding Scheme, SAC Disease Surveillance Centre, Drummond Hill, Inverness IV2 4JZ, UK, 7Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK, 8GECC (Groupe d’Etude des Cétacés du Cotentin), Place des Justes, 50130 Cherbourg-Octeville, France, 9School of Marine Science and Engineering, University of Plymouth, Drake Circus, Plymouth, Devon, UK, 10Universidad del Mar, Instituto de Genética, Oaxaca, México, 11Danish National High-throughput DNA Sequencing Centre, University of Copenhagen, Øster Farimagsgade 2D, 1353 Copenhagen K, Denmark, 12Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden, 13CMPG, Institute of Ecology and Evolution, University of Bern, Bern CH-3012, Switzerland

*These authors contributed equally to this work.

Keywords: SNP discovery, Next-Generation-Sequencing, population genomic resources, filtering pipeline, cetaceans

Correspondence: Marie Louis, Fax: +335-46-44-99-45;
E-mail: marielouis17@hotmail.com
Andrew D. Foote,
E-mail: footead@gmail.com

**Running title:** Bottlenose dolphin SNP discovery

**Abstract**

The discovery and development of a high density, genome-wide SNP array can be an important step towards a thorough understanding of local adaptation, mutation load and demographic history and can therefore inform the conservation and management of species. Bottlenose dolphins, *Tursiops truncatus*, are protected by the Habitats Directive in European waters where they form large offshore and small localized coastal populations. Here, we present a simple and relatively low-cost approach to SNP discovery by shotgun sequencing in this species. We shotgun-sequenced indexed libraries of 33 individuals from the North-East Atlantic. As the mean coverage across sites was low (mean coverage of 4×), we identified SNPs by estimating genotype likelihoods. We then applied a series of stringent filtering steps to remove SNPs in potential repeat regions, paralogous regions, NUMTs and regions of low mappability. This resulted in the discovery of a total of 440,718 SNPs. We validated 266,187 SNPs by comparing our data to two published high coverage (>30×) full genomes of bottlenose dolphins, one originating from the Pacific, and the other from the North-West Atlantic. Whilst this approach only results in SNP discovery rather than simultaneous SNP discovery and genotyping, it only needs to be done once for a set of populations, allowing subsequent studies to target sequence this reference set of SNPs using hybridization-enrichment capture for a broad range of applications (i.e. inferences of selection and demographic history). We highlight how this resource could be used to optimize the number of SNPs targeted in different RAD-seq strategies by simulating RAD-seq experiments with *in silico* endonuclease cutting sites.
Introduction

The advent of high-throughput sequencing technologies coupled with the development of population genomic approaches have facilitated the use of genome-wide data for the inference of local adaptation, historical demography and admixture. In particular, RAD-sequencing (RAD-seq), genotyping-by-sequencing (GBS), and associated reduced representation library (RRL) methods (Baird et al. 2008; Davey et al. 2011; Elshire et al. 2011; Davey et al. 2013; Andrews et al. 2016) allow relatively low-cost simultaneous SNP (Single Nucleotide Polymorphism) discovery and genotyping of large numbers of individuals, and have therefore become increasingly more widely used in empirical studies (see Narum et al. 2013 for a review). RAD sequencing has generated a large number of SNPs in some species such as the European eel (Anguilla anguilla) (376,918 SNPs, Pujolar et al. 2013) and rainbow trout (Oncorhynchus mykiss) (145,168 SNPs, Palti et al. 2014). However, these reduced representations of the genome might not be well suited for some applications for species with low genetic diversity without optimization to increase the number of SNPs targeted. As an example, cetaceans typically have slow molecular clocks (Bininda-Emonds 2007; Jackson et al. 2009; Dornburg et al. 2012; McGowen et al. 2012) and low nucleotide diversity (Table 1, Yim et al. 2014). Consequently a relatively low number of SNPs (typically <10,000) have been discovered by RAD-sequencing studies in this taxonomic group (Moura et al. 2014a; Viricel et al. 2014; Cammen et al. 2015; but see Fernández et al. 2016, Table 2). Given the low genetic diversity in many cetacean species, RRL data may not be ideal for some inferences of intra-specific population history. For example, the accuracy of demographic inference based on the site frequency spectrum (e.g. Gutenkunst et al. 2009; Excoffier et al. 2013; Liu & Fu 2015) is a function of the number of segregating sites (Terhorst & Song 2015). Furthermore, many inferences of intraspecific history including demographic history and selection utilize information from multiple linked polymorphic sites spanning longer contiguous sequences (e.g. Nielsen et al. 2005; Li & Durbin 2011). Therefore, for those particular applications in species with low genetic diversity yielding low density of SNPs within RRLs, either an alternative approach and/or optimization of the RRL method to maximize SNP yield may be needed.
Here, we use an alternative, simple and low cost approach to SNP discovery by shotgun sequencing (SDBSS) in the bottlenose dolphin (Tursiops truncatus), a species with relatively low genetic diversity (Table 1, Yim et al. 2014). Previous studies have identified 153 SNPs by sequencing targeted regions of the genome and 7,431 SNPs using RAD-sequencing in this species (Vollmer & Rosel 2012; Cammen et al. 2015, Table 2). Bottlenose dolphins are long-lived, social marine mammals that have a worldwide distribution. Their range includes a large variety of habitats including temperate and tropical, coastal, deep pelagic and insular waters. Their ecology and morphology are highly variable across their range, with two ecotypes “pelagic” and “coastal” reported in the North-West Atlantic (NWA), North-East Pacific (NEP) and North-East Atlantic (NEA) (Hoelzel et al. 1998; Segura et al. 2006; Louis et al. 2014a; Louis et al. 2014b; Lowther-Thieleking et al. 2015). In the NEA, coastal populations have likely been founded by the pelagic population relatively recently, after the Last Glacial Maxima (10,320 yrBP, 95% CI: 4,300–47,800, Louis et al. 2014a). The most likely hypothesis is that they originated from pelagic individuals that colonized European coastal waters when the sea ice retreated. Therefore, they are an interesting study system to investigate the influence of both adaptive and demographic processes during genomic divergence. As detailed previously, these analyses require the development of a large set of genetic markers. In addition, bottlenose dolphins are protected in European waters by the Habitats Directive (Council Directive 92/43/EEC) and the conservation of the species requires a thorough understanding of local adaptations, population structure, demographic history and gene flow. Efforts to compare among genetic studies focusing on localized populations have so far been hindered by the use of different microsatellite markers. The development of a SNP resource will overcome these comparison issues, allowing future studies to target the same markers and provide a complete understanding of population structure. Moreover, studies on this species can benefit from the availability of two full high coverage (>35×) genomes: a US Navy dolphin from the Gulf of Mexico (Sam Ridgway, personal communication); and a dolphin from the Pacific (Lindblad-Toh et al. 2011; Yim et al. 2014; Foote et al. 2015). Given the recent divergence of bottlenose dolphins in the NEA (Moura et al. 2013; Louis et al. 2014a) and the slow molecular clock of cetaceans (Bininda-Emonds 2007; Jackson et al. 2009; Dornburg et al. 2012; McGowen et al. 2012), many polymorphic sites are likely to be standing genetic variants shared globally across the populations throughout the species’ range and these
high coverage genomes are therefore anticipated to be able to validate many of our inferred SNPs.

For this study, we mined whole genome shotgun-sequencing data that included both coastal and pelagic bottlenose dolphins from the NEA, to investigate the potential for such data to be used for the discovery of SNPs in the nuclear genome. We applied stringent filtering steps to limit errors linked to sequencing and low-coverage data. In order to identify putative global variants, we compared our data to the two available bottlenose dolphin full genomes. In addition, we checked our data against the two full genomes of another delphinid species, the killer whale (*Orcinus orca*) (Moura *et al.* 2014b; Foote *et al.* 2015), to identify SNPs that correspond to potential ancient standing variants within the Delphinidae. We provide a large SNP resource that could be genotyped using hybridization-enrichment capture methods for downstream population genomics analyses of NEA bottlenose dolphins. Users could also use this SNP dataset to customise RAD-seq strategies to optimize the number of targeted SNPs.

**Materials and methods**

*DNA extraction, library preparation and sequencing*

Thirty-three samples of epidermal tissue samples were collected from stranded and biopsied sampled free-ranging individuals (see Louis *et al.* 2014b; Nykänen *et al.* *in preparation*) from France (*N* = 12), England (*N* = 2), Ireland (*N* = 9), Wales (*N* = 3) and Scotland (*N* = 7, Figure 1 plotted using the MARMAP package (Pante & Simon-Bouhet 2013) in R 3.2.0 (R Core Team 2015)). The samples were previously genotyped using microsatellites and genetically assigned (apart from 1 individual) to coastal (*N* = 21) or putative pelagic (*N* = 11) populations using population genetics methods such as STRUCTURE, TESS and DAPC (Ilas 2010; Mirimin *et al.* 2011; Louis *et al.* 2014b). DNA was extracted from epidermis tissue using a Qiagen DNeasy kit following the manufacturer’s guidelines. Genomic DNA was then sheared to an average size of ~150-200 bp using a Diagenode Bioruptor NGS run with 20 cycles of 30 seconds on, and 30 seconds off. Illumina sequencing libraries were built on the sheared DNA extracts using NEBNext (Ipswich, MA, USA) DNA Sample Prep Master Mix Set 1 following Meyer and Kircher (2010). Libraries were subsequently index
amplified for 15 cycles using Phusion High-Fidelity Master Mix (Finnzymes) in 50-μL reactions following the manufacturer guidelines. The libraries were then purified using a MinElute PCR purification kit (Qiagen, Hilden, Germany). The DNA concentration of the libraries was measured using a 2100 Bioanalyzer (Agilent Technologies, CA, USA), these were pooled approximately equimolarly and then sequenced on an Illumina HiSeq 2500 platform as a lane of 100-bp single read (SR) using v4 chemistry.

**Base calling, sequence read trimming and mapping**

Conversion of Illumina's *.bcl files to fastq, and demultiplexing were performed using Illumina’s CASAVA 1.8.2 software allowing for no mismatch in the 6-nucleotide indices used for barcoding. Sequencing reads within the generated fastq files were processed with ADAPTER-REMOVAL (Lindgreen 2012) to trim residual adapter sequence contamination and to remove adapter dimer sequences as well as low-quality stretches at 3’ ends (i.e. consecutive stretches of N’s and of bases with a quality score of 2 or lower). Sequence reads that were ≤30 bp following trimming were discarded. The remaining filtered reads were first mapped to a bottlenose dolphin mitochondrial genome to be removed and analyzed separately. Reads that did not map to the mitochondrial genome were then extracted from the bam file and converted into a fastq file using SAMTOOLS (Li et al. 2009). These reads were then mapped, requiring a mapping quality greater than 30, to the reference bottlenose dolphin genome assembly (Ttru_1.4/turTru2, GenBank Assembly ID: GCA_000151865.2, Lindblad-Toh et al. 2011; Foote et al. 2015) using BWA (v. 0.6.1) (Li & Durbin 2009), which had been hard-masked using REPEATMASKER (Smit et al. 1996) and TANDEM REPEATS FINDER (Benson 1999) and was accessed from the UCSC genome browser (Karolchik et al. 2014). Clonal reads were collapsed using the rmdup function of the SAMTOOLS (v. 1.2.1) suite (Li et al. 2009). Short read data from two bottlenose dolphins and two killer whales *Orcinus orca* (PRJNA20367, PRJNA167475, SRR940825, Lindblad-Toh et al. 2011; Moura et al. 2014b; Yim et al. 2014; Foote et al. 2015) were additionally accessed from the National Center for Biotechnology Information Sequence Read Archive database and mapped to the reference genome assembly as above. Coverage was then estimated for each genome using the doDepth function in the ANGSD software package (Korneliussen et al. 2014).
Multi-sample genotype likelihood calling

Our SNP discovery pipeline followed the suggested best practices outlined in Nielsen et al. (2011). Briefly, we used a multi-sample SNP calling approach, taking uncertainty into account by calculating genotype likelihoods, as recommended by Nielsen et al. (2011). Uncertainty in SNP calling in low coverage data (i.e. sequencing depth <6×) can arise from sequencing, base-calling, mapping and alignment errors. To limit these biases, the quality scores of the sequencing data can be integrated in probabilistic methods to calculate genotype likelihoods. The genotype likelihood is the marginal probability of the read data given the genotype of a particular individual at a particular site that is rescaled by the quality score of each read (Nielsen et al. 2011). In addition, variant discovery accuracy and efficiency with low coverage data are largely improved using multiple samples in comparison with a single sample (Nielsen et al. 2011). First, the reads may have only been sampled on one of the two chromosomes of a diploid individual. Thus, calling SNPs from only one sample or after combining the results from each individual separately would lead to very low power (Li 2011). In addition, multiple samples allow the discovery of SNPs based on the estimated allele frequencies (Nielsen et al. 2011). Sites were called as SNPs if the minor allele frequency was significantly different from 0 as inferred from a likelihood ratio test using a chi-square distribution with one degree of freedom (Kim et al. 2011). We set a conservative threshold of only calling SNPs at sites inferred to be variable with a probability of $P < 0.000001$ by the likelihood ratio test. We calculated the genotype likelihood using the SAMTOOLS method implemented in ANGSD (Li 2011; Korneliussen et al. 2014) based on an Expectation Maximization algorithm to both infer the major and minor alleles and to estimate major and minor allele frequencies (Kim et al. 2011; Skotte et al. 2012). The major allele was inferred and uncertainty in the determination of the minor allele was taken into account by summing over the three possible alleles weighted by their probability (Kim et al. 2011). The identified SNPs could be due to heterozygous sites within individuals, or alternative alleles being sequenced in different individuals. SNPs were called separately for the NEA bottlenose dolphin shotgun sequencing (SGS) data, the two high coverage bottlenose dolphin genomes and the two high coverage killer whales.
SNPs filtering

Five filtering steps were then applied to the SNPs inferred from the genotype likelihood estimation using R 3.2.0 (R Core Team 2015) to further avoid artifacts linked to NGS and low-coverage data. We first filtered SNPs in regions of poor mapping quality (Q<30). Then, we discarded SNPs with a depth of coverage higher than twice the mean coverage (mean coverage is 5×). High coverage of these SNPs is potentially the result of unmasked repeated regions, in nuclear mitochondrial DNA (NUMTs), or some other mapping artifact (e.g. paralogous loci). Regions of poor mapping quality (Q<30) and excessive coverage (>10×) were detected using the CALLABLELOCI tool in GATK (McKenna et al. 2010; DePristo et al. 2011). As a further step to remove regions of excessive coverage which could have arisen due to the mapping issues we highlighted above, we plotted the number of SNPs against the number of individuals and discarded the SNPs in the upper tail of the distributions: specifically SNPs found in more than 10 individuals. As it can be difficult to accurately call variants with MAF < 0.1 (Maruki & Lynch 2015), we further discarded the SNPs with estimated minor allele frequency (MAF) < 0.1. Nevertheless, estimations of genotype likelihoods rather than genotype calling are expected to substantially reduce the error rate. In addition, rare variants are useful for many applications such as for the inference of demographic history based on the SFS and the estimation of several population genetic parameters such as diversity and $F_{ST}$ estimates (Nielsen 2004; Clark et al. 2005). Therefore, the data from this filtering step are retained in the dryad depository.

Each of these four filters was applied to the original SNP set inferred from the genotype likelihood estimates to evaluate the number of kept and discarded SNPs at each step. The extent of overlap in the SNPs removed by the different filtering steps was visualized using a Venn diagram. The objective of this study was to develop a SNP array for NEA bottlenose dolphins, thus, all samples were from the NEA. Therefore, our SNP dataset could be subject to ascertainment bias if used for some population genetics inferences on samples from another geographical area (Nielsen 2004). To define a set of putative global variants for this species, we identified SNPs that were also polymorphic or had alternative alleles in either one or both high coverage sequences of individuals originating from the Pacific and the West Atlantic. All the filtering steps were then applied to the original SNP set to get a final set of high
confidence global variants. Lastly, to identify ancestral polymorphisms within the Delphinidae, we repeated the above, identifying sites that were polymorphic across the bottlenose dolphins and the killer whales.

Population genomic exploration of SNP efficacy

NGSADMIX (Skotte et al. 2013) was used to estimate individual’s ancestry based on genotype likelihoods of the filtered total NEA SNPs (i.e. SNPs that passed the first four filtering steps but that were not filtered based on shared polymorphism with the two high coverage genomes), therefore avoiding inferring individual genotypes and taking the uncertainty in genotype calling into account. NGSADMIX is a maximum-likelihood based clustering method that can provide reliable population structure results with very low coverage data as shown by simulations and real data analyses (Skotte et al. 2013). We acknowledge that our dataset is sparse and contains missing data, and thus population structure inferences should be only considered as exploratory, with the aim of providing some indication that the SNP array is informative (i.e. can detect population structure). In addition, NGSADMIX assumes that the loci are at Hardy-Weinberg Equilibrium (HWE) and we acknowledge that given the coverage of the data we could not test for this. NGSADMIX was run for a number of ancestral populations with K set from 2 to 4 and including all SNPs, polymorphic sites found at sites covered in at least 5 individuals and at least 9 individuals with 3 replicates runs. Individuals’ ancestry proportions were compared to previous microsatellite based studies (Islas 2010; Mirimin et al. 2011; Louis et al. 2014b). The best number of ancestral populations K was inferred based on the known population structure inferred in these microsatellite studies.

In silico endonuclease digestion experiments to customize and optimize RAD-seq strategies

We performed in silico cut experiments to evaluate the number of SNPs generated by different RAD-sequencing strategies using a script which we have provided in the dryad depository. Using the bottlenose dolphin reference genome, RAD data were simulated by sampling sequences proximal to the restriction sites. Different parameters were tested including the enzymes used (NotI, SbfI and a combination of NotI and SbfI, i.e. double digest RAD-seq) and the fragment size selection (minimum and maximum
length of the fragment). We then evaluated the number of total filtered NEA SNPs generated by each RAD-seq strategy.

Estimates of the scaled mutation rate, \( \theta \) from whole genome sequences of several cetacean species

Whole genome sequence reads of single individuals of several cetacean species from Yim et al. (2014), Foote et al. (2015) and Keane et al. (2015) were extracted from the National Center for Biotechnology Information Sequence Read Archive database and assembled as detailed previously. Maximum likelihood estimates of \( \theta \) were computed from the assembled reads using mlRho version 2.8 (Haubold et al. 2010, 2014).

Results

Estimates of the scale mutation rate, \( \theta \) from whole genome sequences of several cetacean species

Diversity estimates were relatively low in all cetacean species (Table 1). Bottlenose dolphins showed intermediate \( \theta \) estimates in comparison to other cetacean species (Table 1).

Raw data

We generated 292×10^6 sequencing reads from our shotgun sequencing data of 33 individuals, of which 205×10^6 uniquely mapped to the reference genome. We retained bases with a minimum Phred score of Q30 (Figure 2a illustrates the distribution of the quality scores). The mapped data from all individuals resulted in a mean of 5× coverage of the genome, but the sequencing reads for each individual covered only a fraction of the genome, with few bases being covered >1× by reads from the same individual (Figure 2b). For 28 of the 33 NEA individuals, >1 million reads were generated for each individual (Supplementary Table 2).

Identification of high likelihood polymorphic sites

The likelihood ratio test inferred 530,844 sites to be polymorphic at a probability of \( P < 0.000001 \) in the NEA bottlenose dolphin SGS data. We further identified 4,466,188 and 1,022,488 SNPs in the two high coverage bottlenose dolphin genomes and the two high coverage killer whale genomes, respectively. After filtering out SNPs (Figure 3a) with poor mapping quality (Q<30), sites with excessive coverage (>10×, Figure 3b),
SNPs sequenced in >10 individuals (Figure 3c), SNPs with a MAF < 0.1 (Figure 3d) and SNPs that were not polymorphic in the two dolphin high coverage genomes, 266,187 global variants were retained. A further set of 174,531 SNPs that passed the first four filtering steps but were only found in the NEA dolphins could be useful for population genomic studies in the NEA region. They may however be prone to ascertainment bias when genotyped in other geographical areas (see discussion). The number of SNPs that were kept and removed by each filtering step is given in Table 3 and lists of the positions of SNPs that were kept after each filtering step are provided in dryad. Relatively high numbers of SNPs were simultaneously removed by the coverage, number of individuals and/or mapping quality filters suggesting that they were unmasked repeats or some other artifact (Figure 4). A total of 33,489 SNPs were shared between the filtered NEA dolphin variants and the two high coverage killer whale sequences, and 169,016 SNPs were also polymorphic across the two high coverage bottlenose dolphins and the two killer whale sequences. 30,932 SNPs were shared between the three datasets. SNP sharing is visualized in a Venn diagram drawn using the online tool available at http://bioinformatics.psb.ugent.be/webtools/Venn/ (Figure 5). The filtered total NEA SNPs (i.e. SNPs that passed the first four filtering steps but were not filtered based on shared polymorphism with the two high coverage genomes) were widely and densely distributed across the genome (Figure 6).

Among the SNPs found in the two high coverage dolphin genomes that had a genotype likelihood higher or equal to 0.99 for both individuals, 66% were heterozygous in one individual and homozygous in the other, 10% were heterozygous within both individuals and 24% were homozygous but with different alleles fixed within each individual. When considering the variants that were also found in the NEA dolphins, these proportions change slightly. 53% of the SNPs were heterozygous in one individual and homozygous in the other, 21% were heterozygous within both individuals and 26% were homozygous within both individuals. The mean number of individuals covered in the NEA shotgun dataset is slightly higher (7.3) for the SNPs that are heterozygous in both high coverage individuals when compared with the SNPs that are heterozygous in one individual and homozygous in the other (6.5) and homozygous in both individuals (6.4). Thus, it may be that some of these sites that are heterozygous in both high coverage individuals represent unmasked paralogues and
we caution against including these in a SNP-typing array. The list of SNPs after removal of these possible unmasked paralogues is provided in dryad.

Population genomic exploration of SNP efficacy

Using SNPs found at sites covered in at least nine individuals (corresponding to 17,866 SNPs) and assuming \( K=2 \), the inferred ancestry proportions were highly consistent with previous \( K=2 \) STRUCTURE runs on the same individuals using microsatellite genotypes (Izas 2010; Mirimin et al. 2011; Louis et al. 2014b). When including SNPs covered in less than nine individuals and for higher values of \( K \), results became less consistent with the previous microsatellite studies. The dolphins were successfully assigned to the coastal and pelagic ecotypes (as determined previously in other studies) with high ancestry proportions apart from one individual (Figure 7). The dataset was however unable to identify the fine-scale population structure found within the coastal ecotype using microsatellite data (Louis et al. 2014b). Bayesian clustering analyses on microsatellite data showed that individuals sampled in the United Kingdom and Ireland formed a separate population (referred to as the “Coastal North” population) from individuals sampled in the English Channel (named the “Coastal South” population). While individuals from the coastal North population were assigned with high ancestry proportions to the coastal ecotype using the SNP dataset, all the individuals from the coastal South population showed some degree of admixture. Including SNPs with MAF < 0.1 or restricting the analysis to the SNPs only found in NEA dolphins did not improve the inference of fine-scale population structure within the coastal ecotype (data not shown). Given that the SNPs used for this inference were only covered (potentially to only 1×) in nine out of the 33 individuals, we stress that this analysis is for exploration only and some affirmation of the potential for downstream use of these SNPs. Nevertheless, given the relative concordance between the result here and previous results using microsatellite genotypes, we anticipate that SNP-typing of these markers to high coverage for all individuals would provide unprecedented resolution of population structure in this species in the NEA.

In silico cut experiments to customize and optimize RAD-seq strategies

The number of total filtered NEA SNPs recovered varies depending on the chosen RAD-seq strategy (Supplementary Table 1). As an example, the maximum number of recovered SNPs was obtained using \( SbfI \) enzyme and no selection on the maximum
size of the RAD fragments. For all the strategies, a relatively low number (<3,000) of SNPs were recovered, validating our assertion that RRLs may not be appropriate for some applications for species with low genetic diversity. Nevertheless, given the fragmented nature of the bottlenose dolphin reference genome that is not yet assembled into chromosomes, the numbers of SNPs within selected fragments are likely underestimated.

Discussion

This study highlights that with careful filtering, shotgun-sequencing data can be opportunistically used for SNP discovery. Low coverage sequencing data are prone to sequencing, base-calling and mapping artifacts that might lead to false-positive polymorphism detection. The use of genotype likelihood methods that take into account the uncertainty in the data, combined with additional validation and stringent filtering steps can overcome these issues and provide high confidence in variant discovery. Several outcomes of our study support this expectation. First, there is relatively large overlap in the SNPs filtered out by each of our steps to remove variants in unmasked repeated regions, NUMTs and regions of low mappability. In addition, we found a large number of polymorphic sites (266,187 SNPs) in our dataset, which were also polymorphic in the two high coverage individuals, validating and giving high confidence to our set of SNPs. This large proportion of shared SNPs between the bottlenose dolphins from different oceans was expected given the relatively recent time to the most recent common ancestor (TMRCA) and slow molecular clocks of cetaceans (Bininda-Emonds 2007; Dornburg et al. 2012; McGowen et al. 2012; Moura et al. 2013). Thus, many polymorphic sites are likely to be standing genetic variation shared globally among populations, rather than derived mutations. This approach could be applied to any species for which a related species reference genome is available to map the reads.

An additional set of 174,531 SNPs that were found in the NEA dolphins were not identified as shared variants with the two high coverage individuals. Whilst false-discovery remains a possibility for these SNPs, our conservative pipeline for identifying SNPs using genotype-likelihoods combined with our stringent filtering steps will have minimized the false-discovery rate. These 174,531 SNPs may be
derived within the NEA populations and may therefore be subject to ascertainment bias if typed in populations that are not from the NEA, but are suitable for population genomic analyses specific to the NEA region. Nevertheless, some of these SNPs could be global variants that are simply by chance homozygous in the two individuals sequenced to high coverage. Ascertainment bias can arise when a small panel of individuals is used to discover SNPs or when SNPs are not geographically representative because they were discovered in one population (Morin et al. 2004; Nielsen 2004). In these conditions, SNPs with low minor frequency alleles are less likely to be discovered than SNPs with intermediate allele frequencies. The Site-Frequency-Spectrum (SFS) in the larger sample of individuals that is typed after the SNP discovery will then be skewed towards an excess of common alleles (Nielsen 2004; Lachance & Tishkoff 2013). Thus, inferences based on the SFS, in particular demographic history, linkage disequilibrium, diversity and \( F_{ST} \) estimates will be biased (Nielsen 2004; Clark et al. 2005; Albrechtsen et al. 2010).

For population genomics analyses on NEA bottlenose dolphins, our SNP array (440,718 total variants) has likely very low ascertainment bias given that 33 individuals from various populations (the Atlantic pelagic population and several coastal populations from the United-Kingdom, Ireland and France, see Louis et al. 2014a,b; Mirimin et al. 2011) have been used in the discovery panel. In addition, the inclusion of the rare variants discovered prior to applying the cut-off at MAF < 0.1 should allow for robust inference of most population genomic parameters in Northeast Atlantic populations. Discarding singletons and rare alleles may lead to bias when inferring demographic history based on the SFS (as they provide a signature of recent population size expansion) and population genetics parameters such as diversity and \( F_{ST} \) estimates (Nielsen 2004; Clark et al. 2005). Therefore, in the dryad folder, we have also provided the list of SNPs that have passed the three first filtering steps, but are not filtered based on MAF (i.e. 453,524 SNPs). We recommend that this set of SNPs is used for population genomics inferences based on the SFS or where singletons and SNPs with low MAFs are important. For population genomics studies in other areas of the distribution range of the species, we acknowledge that the global variants (266,187 SNPs) might not be optimal for some analyses that make inferences from rare variants (such as analyses that depend upon SFS-based inference). However, alternative approaches could be used such as inference of demographic history using
haplotype length (Harris & Nielsen 2013), inferences of selection on standing genetic variation (Seehausen et al. 2014) and test of selective sweeps based on allele frequency differentiation across populations at contiguous multiple loci (Chen et al. 2010). In addition, the SFS and uncertainty in the associated parameter estimates can be corrected for ascertainment bias linked to the fact that low frequency alleles have been discarded (Nielsen et al. 2004). The 171,573 SNPs that are ancestral polymorphisms shared between the killer whale and the bottlenose dolphin could be useful for phylogenetic studies on delphinids. Homoplasy may be an underlying process for some of these SNPs, meaning that some of these SNPs might not be broadly found in other delphinids. However, incomplete lineage sorting may also led to shared variants (as suggested for great apes, Pruefer et al. 2012; Mailund et al. 2014). A previous study (Fernández et al. 2016) found a high proportion of shared SNPs between white-beaked (Lagenorhynchus albirostris) and Atlantic white-sided dolphins (L. acutus) also consistent with incomplete lineage sorting.

The coverage of our data was too low and variable to infer individuals’ genotypes and perform population genomic analyses that incorporated all discovered SNPs, even with methods taking uncertainty into account (Li 2011; Nielsen et al. 2012; Maruki & Lynch 2015). Although a clustering algorithm based on genotype likelihoods (Skotte et al. 2013) reliably assigned the dolphins to the coastal and pelagic ecotypes, the dataset is too incomplete to detect the finer-scale population structure within the coastal ecotype that was found using microsatellite data (Louis et al. 2014b). Nevertheless, this result strongly suggests that our SNP array is informative. Thus, in comparison to RRL methods such as RAD-seq and GBS, SNP discovery by shotgun sequencing (SDBSS) resulted only in SNP discovery rather than simultaneous SNP discovery and genotyping. Nevertheless, SDBSS potentially only needs to be done once for a species allowing future studies to target-sequence this reference set of SNPs using high-density SNP arrays or custom-produced baits for enrichment capture. Briefly, DNA libraries would be enriched with custom-designed biotinylated RNA baits through a hybridization reaction to capture targeted loci with sequences that are identical to the set of baits (Gnirke et al. 2009). Here, baits could be produced using the SNPs’ coordinates on the bottlenose dolphin reference genome. For applications in species for which a reference genome is not available, the genome of a related species can be used. Enriched libraries would then be PCR-amplified and sequenced
using a next-generation technology. Currently custom baits can be used to genotype hundreds of thousands of SNPs (up to 200,000 loci) for tens to thousands of individuals. The main advantage of this approach is to target and sequence only loci that are polymorphic, which is likely to be cost effective for species with low genetic diversity such as cetaceans. For example, a recent RAD-seq study on two dolphin species by Fernández et al. (2016) found that 68.3% of RAD-tags were monomorphic. This set of SNPs could be a basis for future population genomics studies on bottlenose dolphins in the NEA, and foster population structure studies that would be comparable between laboratories/geographical regions. The bottlenose dolphin is a species of strong conservation focus, as it is one of just two cetacean species listed on the European Habitats Directive. In addition, coastal populations are small, relatively isolated and have restricted home ranges which raise conservation concerns in the context of global changes (Ingram & Rogan 2002; Mirimin et al. 2011; Berrow et al. 2012; Cheney et al. 2014; Louis et al. 2014b; Louis et al. 2015). However, previous population genetic studies in the NEA have used microsatellites, which can constrain comparison between laboratories and datasets giving a fragmented picture of population structure. In addition, targeting-sequencing these set of SNPs using hybridization capture would also allow the investigation of local adaptation (e.g. Chen et al. 2010), mutation load (e.g. Peischl et al. 2013) and demographic history (e.g. Excoffier et al. 2013; Liu & Fu 2015), all of which are likely to play a role in the conservation of the recently founded and small coastal populations (Louis et al. 2014a,b). As stated earlier, these analyses require a large number of loci (Terhorst & Song 2015) or multiple linked variants (Chen et al. 2010) such as provided here.

Our shotgun approach can also be coupled with RAD-seq to optimize genotyping-by-sequencing efforts. Recently, in silico cut experiments have been used to estimate the expected number of SNPs under different RAD-seq strategies (DaCosta & Sorenson 2014; Lepais & Weir 2014). But this customization may not be trivial without any reference set upon which to base the optimization. We highlighted how RAD-seq strategy could be optimized by choosing the enzyme and size selection that are suitable for a given number of SNPs (Supplementary Table 1). We acknowledge that the number of SNPs recovered by the in silico cuts are likely underestimated due the fragmented nature of the bottlenose dolphin genome. Nonetheless, these results do provide a measure of the relative number of SNPs expected to be sequenced by
different RAD strategies, and also highlight that RAD-seq is likely to result in the
discovery of orders of magnitude fewer SNPs than identified by the shotgun
sequencing approach applied here. The RAD-seq customization approach could be
suitable for downstream analyses for which fewer SNPs are needed, e.g. sequencing
large number of individuals to investigate population structure.

Conclusions

This methodology, in particular the validation and filtering steps could be a template
for SNP discovery by shotgun sequencing (SDBSS) based on genotype likelihoods for
future studies. In contrast to RRL approaches, SDBSS only discovered SNPs and did
not simultaneously genotype individuals. However, this approach generated a large
SNP resource for NEA bottlenose dolphins that provides scope for a wide range of
population genomics analyses of Northeast Atlantic bottlenose dolphins. We anticipate
population-level sequencing of these SNPs will greatly elucidate the evolutionary
history and provide new conservation insights into locally adaptive genomic changes
in these coastal and pelagic ecotypes of bottlenose dolphins. We also highlight how
this resource could be used to customize and optimize the number of SNPs targeted in
different RAD-seq strategies for delphinids and other species.

Acknowledgements

We would like to thank Pernille Selmer Olsen and Tina Brand for help and advice in
the laboratory; Andaine Seguin-Orlando, Cecilie Mortensen and Lillian Petersen at the
Danish National High-throughput DNA Sequencing Centre for their support with
laboratory analyses. We also thank everyone involved in sample collection: the UK
Cetacean Strandings Investigation Programme, GECC and Réseau National
Echouages volunteers, Fabien Demaret, Ghislain Doremus, Vincent Ridoux and
Olivier Van Canneyt (Observatoire Pelagis), Anneli Englund, Allessandro Ponzo and
Ailbhe Kavanagh, Irish Cetacean Genetic Tissue Bank, Natural History Museum,
Dublin. All samples were taken under the relevant permits. UK samples were collected
under the aegis of the UK Cetacean Strandings Investigation Programme, which is
funded by Defra and the Devolved Governments of Scotland and Wales. The Irish
biopsy samples were collected under permit from National Parks and Wildlife Service,
Department of Environment, Heritage and Local Government in Ireland. The biopsy
samples from France were collected under permit from the Environment Ministry. Funding for sample collection in France was provided by Fondation Total, Agence de l’Eau Seine-Normandie, Fonds de Dotation pour la Biodiversité, Agence des Aires Marines Protégées and Ministère de l’Ecologie, du Développement Durable et de l’Energie, and in Ireland by the Crawford-Hayes studentship fund and by the National Parks and Wildlife Service. ADF was funded by Swiss SNSF Grant 31003A-143393 awarded to Laurent Excoffier.

References


Nykänen M, Louis M, Kaschner K, et al. (in preparation) Reconstructing the post-glacial colonisation of the northern extreme of the range of a top marine predator, the bottlenose dolphin.


Data Accessibility

The following data will be deposited on dryad (please see the readme file for details):

- The fastq files from each NEA bottlenose dolphin
- The original NEA dolphin SNP set inferred from the genotype likelihood estimates
- The NEA dolphin SNPs after each of the filtering steps
- The total NEA dolphin filtered SNPs
- The global filtered SNPs
- The total NEA dolphin SNPs that have passed the three first filtering steps, but are not filtered based on MAF
- The global SNPs that have passed the three first filtering steps, but are not filtered based on MAF
- The total NEA dolphin filtered SNPs with the possible unmasked paralogous loci removed
- The global filtered SNPs with the possible unmasked paralogous loci removed
- The high coverage bottlenose dolphin and high coverage killer whale SNPs inferred from the genotype likelihood estimates
- The SNPs shared between either the two killer whale high coverage sequences and the filtered NEA dolphins SNPs or the two killer whale high coverage sequences and the two dolphin high coverage sequences
- The script to perform the in silico cuts

Author Contributions
ADF designed the study; AB, WD, RD, FG, SI, VI, ML, MN and ER collected the samples; ADF, MN, NW performed laboratory work (DNA extraction, library preparation and sequencing); ADF, SG, ML, KM, BSB and NV conducted the data treatment and analyses; ADF and ML wrote the manuscript and all authors commented on the manuscript.
Table 1. Estimates of the scale mutation rate, $\theta$ from whole genome sequences of several cetacean species using mlRho version 2.8 and nucleotide diversity (mean per-nucleotide heterozygosity) estimated in Yim et al. 2014.

<table>
<thead>
<tr>
<th>Species</th>
<th>$\Theta$</th>
<th>Nucleotide diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottlenose dolphins (<em>Tursiops truncatus</em>)</td>
<td>0.00193</td>
<td>0.00142</td>
</tr>
<tr>
<td>Bowhead whale (<em>Balaena mysticetus</em>)</td>
<td>0.0000158</td>
<td>NA</td>
</tr>
<tr>
<td>Finless porpoise (<em>Neophocaena phocaenoides</em>)</td>
<td>0.0857</td>
<td>0.00086</td>
</tr>
<tr>
<td>Fin whale (<em>Balaenoptera physalus</em>)</td>
<td>0.0569</td>
<td>0.00151</td>
</tr>
<tr>
<td>Killer whale – Atlantic (<em>Orcinus orca</em>)</td>
<td>0.000484</td>
<td>NA</td>
</tr>
<tr>
<td>North Pacific minke whale (<em>Balaenoptera acutorostrata scammoni</em>)</td>
<td>NA</td>
<td>0.00061</td>
</tr>
</tbody>
</table>
Table 2. Number of RAD loci generated and number of SNPs discovered in RAD-sequencing studies on cetaceans

<table>
<thead>
<tr>
<th>Species</th>
<th>Enzyme</th>
<th>No. of loci</th>
<th>No. of SNPs</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>common dolphin (<em>Delphinus delphis</em>) and harbor porpoise (<em>Phocoena phocoena</em>)</td>
<td><em>Not</em>1</td>
<td>5,182</td>
<td>3,595 loci*</td>
<td><em>Viricel et al. 2014</em></td>
</tr>
<tr>
<td>killer whale (<em>Orcinus orca</em>)</td>
<td><em>Not</em>1</td>
<td>Not reported</td>
<td>3,281</td>
<td><em>Moura et al. 2014</em></td>
</tr>
<tr>
<td>white-beaked (<em>Lagenorhynchus albirostris</em>) and Atlantic white-sided dolphins (<em>L. acutus</em>)</td>
<td><em>SbfI-HF</em></td>
<td>179,170</td>
<td>52,981</td>
<td><em>Fernandez et al. 2016</em></td>
</tr>
<tr>
<td>Bottlenose dolphins (<em>Tursiops truncatus</em>)</td>
<td><em>SbfI-HF</em></td>
<td>129,594</td>
<td>7,431</td>
<td><em>Cammen et al. 2015</em></td>
</tr>
</tbody>
</table>

*The number of polymorphic loci and not SNPs are reported.*
Table 3. Number of SNPs that were kept and removed by each filtering step applied to the original NEA dolphin SNP set inferred from the genotype likelihood estimates

<table>
<thead>
<tr>
<th>Filtering step</th>
<th>No. of kept SNPs</th>
<th>No. of removed SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping quality (Q&lt;30)</td>
<td>503,362</td>
<td>27,482</td>
</tr>
<tr>
<td>Coverage (&gt;10×)</td>
<td>520,512</td>
<td>10,332</td>
</tr>
<tr>
<td>Number of individuals (&gt;10)</td>
<td>459,947</td>
<td>70,897</td>
</tr>
<tr>
<td>MAF &lt; 0.1</td>
<td>516,494</td>
<td>14,350</td>
</tr>
<tr>
<td>Not variant in the two high coverage <em>Tursiops</em> genomes</td>
<td>334,733</td>
<td>196,111</td>
</tr>
</tbody>
</table>
Figure legends

Figure 1. Sample locations of coastal and pelagic bottlenose dolphins in the North-East Atlantic.

Figure 2. Quality of the sequenced data: a) distribution of the Phred quality scores and b) counts of sites covered by sequencing reads at different depth for each of the 33 samples. For each sequencing depth value there are 33 bars representing the 33 individuals. These graphs were plotted using the R script available at: https://github.com/mfumagalli/ngsTools/blob/master/scripts/plotQC.R.

Figure 3a. Flow chart of the SNP filtering steps.

Figure 3b. Histogram of the number of sites (i.e. bases) against the sequencing depth in the bottlenose dolphin shotgun sequencing (SGS) data.

Figure 3c. Histogram of the number of SNPs against the number of individuals in the bottlenose dolphin SGS data.

Figure 3d. Histogram of the number of SNPs against the minor allele frequency (MAF) in the bottlenose dolphin SGS data.

Figure 4. Venn diagram of the extent of overlap in the SNPs removed by the different filtering steps. Numbers indicated the amount of SNPs that were excluded by one or several filtering steps. For example, the same 11 SNPs were removed by all filtering steps and the same 388 SNPs were removed by the MAFs and number of individual filters.

Figure 5. Venn diagram of the overlap between the SNPs discovered in the NEA bottlenose dolphin shotgun sequencing data, the two high coverage bottlenose dolphin genomes and the two high coverage killer whale genomes.

Figure 6. Plot of the number of SNPs per scaffold (total variants found in the NEA bottlenose dolphin dataset) as a function of the scaffold length, up to a maximum length of 976,602 bp.

Figure 7. Ancestry proportions of individual bottlenose dolphins inferred using NGSADMIX for K=2 and SNPs found at sites covered in at least 9 individuals and comparison with the ecotype/population inferred using microsatellites in previous studies (Ilas, 2010; Mirimim et al. 2011; Louis et al. 2014b). * indicated the individual that was assigned to the coastal ecotype in previous microsatellites studies.
but was assigned to the pelagic ecotype in the present study. The “unknown” individual is the individual for which no microsatellite data were available.
Figures

Figure 1

Figure 2.a
Figure 2.b

Counts of sites covered by sequencing reads vs. sample depth.

Figure 3.a

- **Step 1**: Remove SNPs of mapping quality <30
  - Detect regions of poor mapping quality (Q<30) using the CallableLoci tool in GATK

- **Step 2**: Remove SNPs covered by >10x reads
  - Discard the SNPs that have a coverage > 2 * mean coverage
  - Detect regions of excessive coverage (>10x) using the CallableLoci tool in GATK

- **Step 3**: Remove SNPs covered by >10 individuals
  - Plot the number of SNPs against the number of individuals
  - Discard the SNPs in the upper tail of the distribution (here >10 individuals)

- **Step 4**: Remove SNPs of Minor Allele Frequency (MAF)<0.10
  - Discard SNPs with estimated MAF<0.1 as it is difficult to distinguish singletons and rare variants from sequencing errors

- **Step 5**: Identify global variants
  - Remove variable sites which were not shared variants with the two high coverage individuals coming from two distinct geographical regions
Figure 3.d

![Bar chart showing minor allele frequency distribution with a peak at 0.5 and a secondary peak at 0.1.]

Figure 4

![Venn diagram illustrating the overlap of sites covered by >10x reads, MAFs <0.1, sites covered in >10 individuals, and mapping quality <30.]

- MAFs <0.1: 12,806
- Sites covered in >10 individuals: 39,329
- Sites covered by >10x reads: 6,224
- Mapping quality <30: 196
- Intersection of all criteria: 4

249
Figure 5

High coverage (>30x) *Tursiops*

<table>
<thead>
<tr>
<th></th>
<th>NEA <em>Tursiops</em> total variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>4,061,917</td>
<td>235,255</td>
</tr>
<tr>
<td>138,084</td>
<td>30,932</td>
</tr>
<tr>
<td></td>
<td>171,974</td>
</tr>
<tr>
<td></td>
<td>850,915</td>
</tr>
</tbody>
</table>

High coverage *Orcinus Orca*

Figure 6

Number of SNPs discovered vs. Length of scaffold (bp)
Supporting Information

High density, genome-wide SNP discovery in Northeast Atlantic bottlenose dolphins based on genotype likelihoods from multiplex shotgun sequencing data

MARIE LOUIS, MILAJA NYKÄNEN, SHYAM GOPALAKRISHNAN, ANDREW BROWNLOW, WILLY DABIN, ROBERT DEAVILLE, EILEEN DILLANE, FRANCOIS GALLY, SIMON N. INGRAM, VALENTINA ISLAS VILLANUEVA, KIM MAGNUSSEN, EMER ROGAN, BENOIT SIMON-BOUHET, NAGARJUN VIJAY, NATHAN WALES and ANDREW D. FOOTE

1Observatoire PELAGIS, UMS 3462 CNRS/Université de La Rochelle, Pôle Analytique, 5 allées de l'Océan, 17000 La Rochelle, France, 2Centre d'Etudes Biologiques de Chizé, CNRS-UMR 7372, 79360 Villiers-en-Bois, France, 3Scottish Oceans Institute, University of St Andrews, St Andrews, Fife, KY16 8LB, Scotland, 4School of Biological, Earth and Environmental Sciences, University College Cork, Ireland, 5Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark, 6Scottish Marine Animal Stranding Scheme, SAC Disease Surveillance Centre, Drummond Hill, Inverness IV2 4JZ, UK, 7Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK, 8GECC (Groupe d’Etude des Cétacés du Cotentin), Place des Justes, 50130 Cherbourg-Octeville, France, 9School of Marine Science and Engineering, University of Plymouth, Drake Circus, Plymouth, Devon, UK, 10Universidad del Mar, Instituto de Genética, Oaxaca, México, 11Danish National High-throughput DNA Sequencing Centre, University of Copenhagen, Øster Farimagsgade 2D, 1353 Copenhagen K, Denmark, 12Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden, 13CMPG, Institute of Ecology and Evolution, University of Bern, Bern CH-3012, Switzerland

*These authors contributed equally to this work.
Supplementary Table 1. Results of the in-silico cut experiments to optimize RAD-seq strategies. In-silico cuts, simulating RAD-seq laboratory protocols, were generated using different enzymes and different fragment size selections. The number of our discovered filtered total NEA dolphin SNPs proximal to cut sites are reported for each RAD-seq strategy.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Total number of cut sites</th>
<th>Total number of fragments</th>
<th>Fragment length Min</th>
<th>Number of fragments passing size selection</th>
<th>Number of SNPs proximal to cut sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>NotI</td>
<td>8084</td>
<td>3468</td>
<td>0</td>
<td>3461</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 1000</td>
<td>699</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 750</td>
<td>638</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 500</td>
<td>531</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 --</td>
<td>3294</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 1000</td>
<td>532</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 750</td>
<td>471</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 500</td>
<td>364</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 --</td>
<td>3216</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>150 1000</td>
<td>454</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>150 750</td>
<td>393</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>150 500</td>
<td>286</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>200 --</td>
<td>3152</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>200 1000</td>
<td>390</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>200 750</td>
<td>329</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>200 500</td>
<td>222</td>
<td>4</td>
</tr>
<tr>
<td>SbfI</td>
<td>52049</td>
<td>35032</td>
<td>0</td>
<td>35032</td>
<td>2831</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 1000</td>
<td>3935</td>
<td>245</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 750</td>
<td>3219</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 500</td>
<td>2464</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 --</td>
<td>34295</td>
<td>2812</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 1000</td>
<td>3198</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 750</td>
<td>2482</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 500</td>
<td>1727</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 --</td>
<td>34060</td>
<td>2806</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>150 1000</td>
<td>2963</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>150 750</td>
<td>2247</td>
<td>148</td>
</tr>
<tr>
<td>NotI-SbfI</td>
<td>60133</td>
<td>8911</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>--</td>
<td>8906</td>
<td>582</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1000</td>
<td>1120</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>750</td>
<td>899</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>500</td>
<td>657</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>--</td>
<td>8718</td>
<td>581</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1000</td>
<td>932</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>750</td>
<td>711</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>500</td>
<td>469</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>--</td>
<td>8633</td>
<td>577</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>1000</td>
<td>847</td>
<td>53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>750</td>
<td>626</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>500</td>
<td>384</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>--</td>
<td>8569</td>
<td>574</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>1000</td>
<td>783</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>750</td>
<td>562</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>500</td>
<td>320</td>
<td>28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplementary table 2. Number of sequence reads generated for each North-East Atlantic individual.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Number of reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>528 549</td>
</tr>
<tr>
<td>51</td>
<td>3 946 390</td>
</tr>
<tr>
<td>33</td>
<td>3 866 286</td>
</tr>
<tr>
<td>101</td>
<td>7 800 637</td>
</tr>
<tr>
<td>7</td>
<td>4 483 754</td>
</tr>
<tr>
<td>A42</td>
<td>210 073</td>
</tr>
<tr>
<td>A41</td>
<td>2 424 750</td>
</tr>
<tr>
<td>B14</td>
<td>2 397 060</td>
</tr>
<tr>
<td>A1</td>
<td>373 430</td>
</tr>
<tr>
<td>A38</td>
<td>4 575 655</td>
</tr>
<tr>
<td>A39</td>
<td>1 252 731</td>
</tr>
<tr>
<td>A4</td>
<td>1 479 072</td>
</tr>
<tr>
<td>SW 2001/141</td>
<td>1 571 916</td>
</tr>
<tr>
<td>SW1991/130</td>
<td>11 749 067</td>
</tr>
<tr>
<td>Tt-09-07</td>
<td>4 517</td>
</tr>
<tr>
<td>Tt-09-10</td>
<td>2 160 320</td>
</tr>
<tr>
<td>Tt-09-01</td>
<td>3 122 802</td>
</tr>
<tr>
<td>Tt-09-04</td>
<td>3 421 701</td>
</tr>
<tr>
<td>Tt-09-03</td>
<td>1 988 400</td>
</tr>
<tr>
<td>Tt-07-01</td>
<td>3 981 043</td>
</tr>
<tr>
<td>Tt-05-14</td>
<td>15 609 814</td>
</tr>
<tr>
<td>2007.1.179</td>
<td>5 196 132</td>
</tr>
<tr>
<td>2007.1.181</td>
<td>14 102 384</td>
</tr>
<tr>
<td>SW 2000/115</td>
<td>6 034 120</td>
</tr>
<tr>
<td>SW 1993/115</td>
<td>19 734 026</td>
</tr>
<tr>
<td>SW 1991/85</td>
<td>6 553 955</td>
</tr>
<tr>
<td>SW 1998/18a</td>
<td>1 258 237</td>
</tr>
<tr>
<td>SW 1997/171b</td>
<td>6 821 424</td>
</tr>
<tr>
<td>SW 2000/141e</td>
<td>12 826 418</td>
</tr>
<tr>
<td>SW 2004/240b</td>
<td>13 327 130</td>
</tr>
<tr>
<td>SW 2000/138a</td>
<td>15 787 409</td>
</tr>
<tr>
<td>SW 2008/93a</td>
<td>199 905</td>
</tr>
<tr>
<td>SW 1994/56f</td>
<td>25 886 537</td>
</tr>
</tbody>
</table>