FINAL REPORT
Focus Group FG21: Developing an approach for integrating genetics into MPA design and spatial planning processes
3rd International Marine Conservation Congress (IMCC3)
19th August, Glasgow, Scotland

Organizers:

Francine Kershaw
Columbia University
Email: flk2107@columbia.edu

Howard C. Rosenbaum, Ph.D.
Wildlife Conservation Society
Email: hrosenbaum@wcs.org

Participants:

C. Scott Baker
Oregon State University
Email: scott.baker@oregonstate.edu

Will McClintock
University of California – Santa Barbara
Email: mcclintock@msi.ucsb.edu

Maria Beger
University of Queensland
Email: m.beger@uq.edu.au

Martín Mendez
Wildlife Conservation Society
Email: mmendez@wcs.org

Kate Blackwell
George Mason University
Email: cblackw4@gmu.edu

Manuel Muntoni
University of Cagliari
Email: m.muntoni@hotmail.it

Inês Carvalho
Instituto Gulbenkian de Ciência
Email: carvalho.inesc@gmail.com

Les Noble
University of Aberdeen
Email: l.r.noble@abdn.ac.uk

Sal Cerchio
Wildlife Conservation Society
Email: scerchio@wcs.org

Delphine Rocklin
Universidad de Murcia
Email: delphine.rocklin@gmail.com

Michelle Frost
University of Aberdeen
Email: r01mbm13@abdn.ac.uk

James Thorburn
University of Aberdeen
Email: j.thorburn@abdn.ac.uk

Oscar Gaggiotti
University of St. Andrews
Email: oeg@st-andrews.ac.uk

Rationale:
The Focus Group addressed the following questions: **How can we integrate genetic information into marine protection and spatial planning tools?** In essence, **How can we make marine science matter?**

The persistence of biodiversity requires the protection of evolutionary processes at the scale of ecosystems, species, and populations. The importance of accounting for evolutionary processes in marine spatial planning efforts has been reflected in a range of policy mechanisms, including the Convention on Biological Diversity (CBD). Genetic tools can provide information that directly responds to this need, including the identification and distribution of species and population "units", delineating habitat use of distinct populations, elucidating patterns of genetic connectivity - essential to the ecological coherence of MPA networks, and estimating effective population sizes. Protection of genetic diversity is becoming increasingly important in the context of environmental disruption, such as climate change. Integration of genetic information into marine spatial planning has so far been lacking, however, despite significant scientific advances in the application of genetics in marine conservation and management.

Goal:
The goal of the Focus Group was to develop an approach to addressing the knowledge and implementation gap currently existing in the marine planning and conservation community regarding the integration of genetic information into marine protected area design and spatial planning efforts. It brought together geneticists and spatial ecologists working in the field of marine protection with technical developers of marine planning tools to establish an approach to integrate genetic information in a way that is both *ecologically meaningful* and *useful for managers*.

Specifically, the following questions were addressed:
1) Why is it essential that genetic information be integrated into marine protection and spatial planning efforts?
2) What are the different types of genetic information that can be useful to marine protection and spatial planning?
3) How can genetic information be transformed so that it can be visualized within a marine planning tool in a way that is both ecologically meaningful and useful for the marine conservation and management community?

Outputs:
The primary output of the proposed Focus Group will comprise a short, illustrated guidance document that will address each of the three questions outlined in the Focus Group's goals. The guidance document will be distributed to a number of international partnerships currently engaged in marine spatial planning efforts, including the Global Ocean Biodiversity Initiative, the Joint IUCN SSC/WCPA Marine Mammal Protected Area Task Force, and IUCN's Key Biodiversity Area team. The Focus Group organizers also intend to present the key findings and recommendations at various fora, to raise awareness of this work both in the scientific and marine conservation and policy communities.

As a secondary output, it is expected that a scientific article outlining the technical aspects of transforming and visualizing genetic data for use in marine spatial planning, and related caveats and considerations, will be prepared for publication in a scientific journal.

***Schedule*:**

9:00 Introduction (Howard Rosenbaum)

9:30 Presentations I: Integrating genetics into MSP - current approaches and tools.
1) Addressing the genetics-policy implementation gap (Francine Kershaw)
2) Marine Spatial Planning 2.0: Integrating genetic and environmental information for wildlife conservation (Martin Mendez)
3) Population genetic tools for quantifying connectivity (Oscar Gaggiotti)

10:15 Break

10:30 Presentations II: Integrating genetics into MSP - current approaches and tools.
4) Genetics as a tool for the spatial management of elasmobranchs (Les Noble, Michelle Frost & James Thorburn)
5) Geodesign using SeaSketch (Will McClintock)
6) GeneGIS - integrating DNA profiles and photo-identification to inform marine spatial planning (C. Scott Baker)

11:15 Brainstorm (2-3 groups): Why is genetic data/information important for MSP?
• What are the data/protection gaps that exist if genetic data isn't included in marine protected area and spatial planning?
• What types of genetic data are useful for MSP? How is it formatted?

12:00 Break

1:00 Break Out Groups: How can genetic data/information be integrated into MSP tools?
• One group for each type of genetic data identified during the Brainstorm session.
• Key questions:
    1. How can genetic data/information be visualized within a MSP tool in a way that improves its <u>ecological meaning</u> and <u>usefulness</u> for the marine conservation and management community?
    2. What are the challenges and potential pitfalls?

2:30 Summary of Group Discussions

3:00 Closing remarks (H. Rosenbaum)

Summary of Presentations:

1) Addressing the genetics-policy implementation gap (Francine Kershaw)

Genetic tools provide unique information for marine spatial planning, protection and management. The genetic toolbox can be used to address questions at multiple scales, including the individual (e.g., parentage analysis), population (e.g., demographic population units, genetic connectivity, effective population size, genetic diversity), and species (e.g., taxonomic units, cryptic species, natural selection). However, while the importance of accounting for evolutionary process in ecosystem management has been recognized by a number of policy mechanisms, the integration of genetic information into marine protected area and spatial planning efforts has been generally lacking.

In 2008, the Convention on Biological Diversity (CBD) agreed on a set of seven scientific criteria to guide the identification and prioritization of Ecologically or Biologically Significant marine Areas (EBSAs): 1) Uniqueness or rarity; 2) Special importance for life history stages of species; 3) Importance for threatened, endangered, or declining species and/or habitats; 4) Vulnerability, fragility, sensitivity, slow recovery; 5) Biological productivity; 6) Biological diversity; and 7) Naturalness. These criteria have subsequently been formally adopted by a number of other initiatives, including the IUCN Joint WCPA/SSC Marine Mammal Protected Area Task Force for the identification of Important Marine Mammal Areas (IMMAs) and by the IUCN for the identification of Key Biodiversity Areas (KBAs). Despite genetic information being informative in the application of each of these criteria, the identification of EBSAs has so far primarily relied upon field surveys, species, censuses, habitat mapping, satellite tracking studies, and expert opinion.

The reasoning behind this implementation gap is not because the integration of genetic information into marine spatial planning is not possible. Indeed, genetic information has been successfully used in a number of marine protected area planning efforts and there are an increasing number of new techniques being developed to spatially represent genetic information. Rather, a knowledge and communication gap between geneticists and the marine policy and management community appears to be the primary driver. To address this gap, a standardized approach to mainstream genetics into marine protection and spatial planning is required. The ability to visualize genetic information alongside other data types commonly used for spatial planning (e.g., habitat models, satellite tracking) in a way that is ecologically meaningful and useful for planners represents an important aspect of developing this approach. Finally, clear guidance to the policy and management community on why this information is important and how to interpret it is essential.
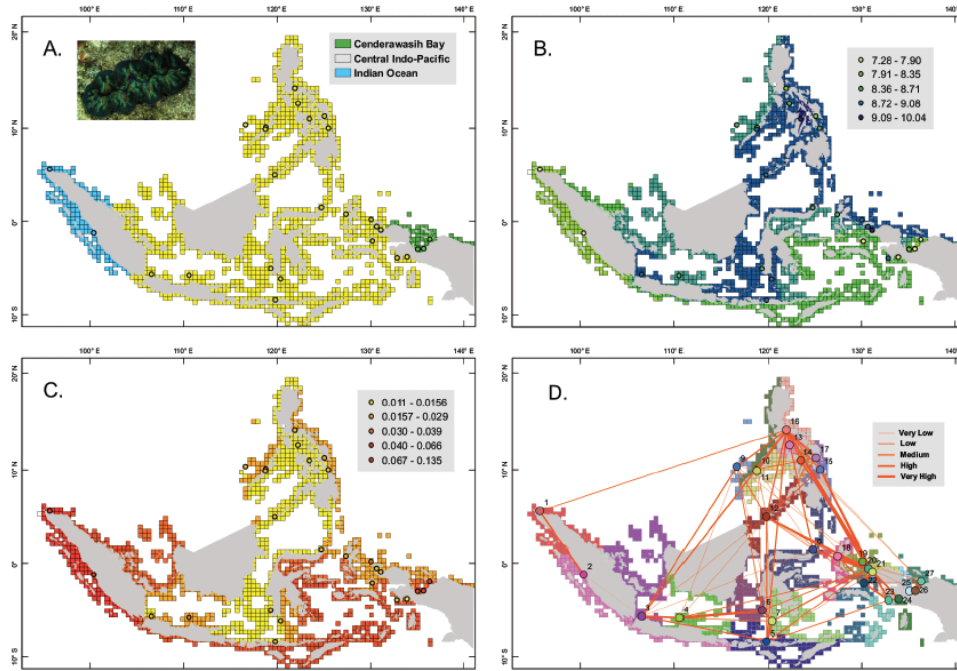
Data visualization

Figure 1: Spatial representation of four genetic attributes: A) Genetic sub-regions (STRUCTURE cluster analysis); B) Genetic diversity (allelic richness); C) Genetic distinctness (local FST); and D) Rates and directionality of gene flow (Beger et al. 2014).
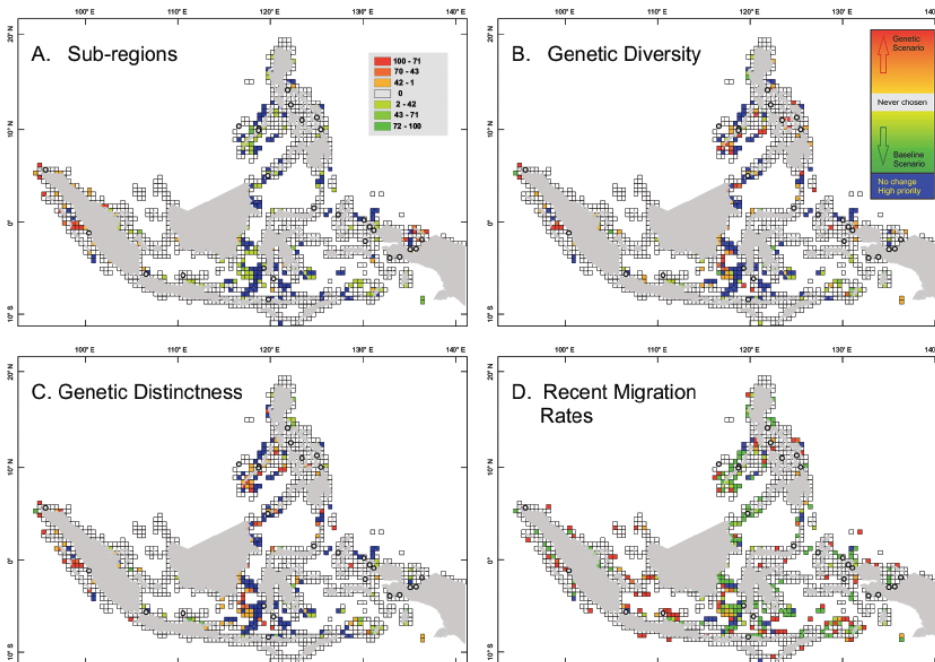

Figure 2: Conservation priority areas identified based on the four genetic attributes in Figure 1 (Beger et al. 2014).

2) Marine Spatial Planning 2.0: Integrating genetic and environmental information for wildlife conservation (Martin Mendez)

Marine spatial planning seeks to conserve the *functional* characteristics of marine areas, accounting for both species and ecological processes. Populations remain the primary unit targeted by conservation strategies, linking evolutionary and ecological scales that ensure long-term species persistence. However, the identification of population units is not trivial, particularly for continuously distributed, highly mobile species. Genetic information can provide evidence for, and quantify, population structure. Integrating genetic information on population structure with spatially explicit information on oceanographic variables (e.g., sea surface temperature, primary productivity, ocean currents) can provide the ecological context of the population units identified, thus implying the ecological drivers of population structure.

Coastal cetaceans represent a useful study system in which to examine the environmental influences on population structure: they are highly mobile species with seemingly continuous distributions, and coastal areas are relatively environmentally heterogeneous. Two studies were conducted on different species in various ocean basins to ensure independent environmental influences: franciscana dolphin (*Pontoporia blainvillei*) off the coast of Argentina and the Indo-Pacific humpback dolphin (*Sousa chinensis*) off East Africa. Genetic tools (mitochondrial DNA and nuclear microsatellites) were used to identify and quantify patterns of population structure and to estimate the magnitude and directionality of their migrations. Ten years of remotely sensed data were used to characterize major oceanographic features in the region that may influence cetacean dispersal (e.g., currents, eddies, environmental boundaries) and the potential differences between the considered environments that the identified populations of dolphins inhabit.

For the franciscana dolphin, boundaries between population units identified by genetic data were found to correlate with environmental boundaries, indicating that oceanographic conditions may be an important process driving population structure (i.e., they are isolated by *environmental* distance, IBED). The utility of this approach for MSP is that regions with different populations units and oceanographic features can be represented visually in the same map. Within Argentina's Exclusive Economic Zone (EEZ), the highest genetic hierarchy corresponded to oceanographic provinces, however subpopulation structure within provinces exists due to demographic history or behavior. MPAs in the region therefore do not currently account for the fact that different genetically distinct populations are subject to differences in threat levels from bycatch. Similar significant correlations between genetic and environmental boundaries were observed for the humpback dolphin, with oceanographic provinces again representing adequate proxies for population isolation. However, scale is an important consideration and examination of this system at the local level is required, given that it is common to see finer-scale population structure indicative of demographic isolation.
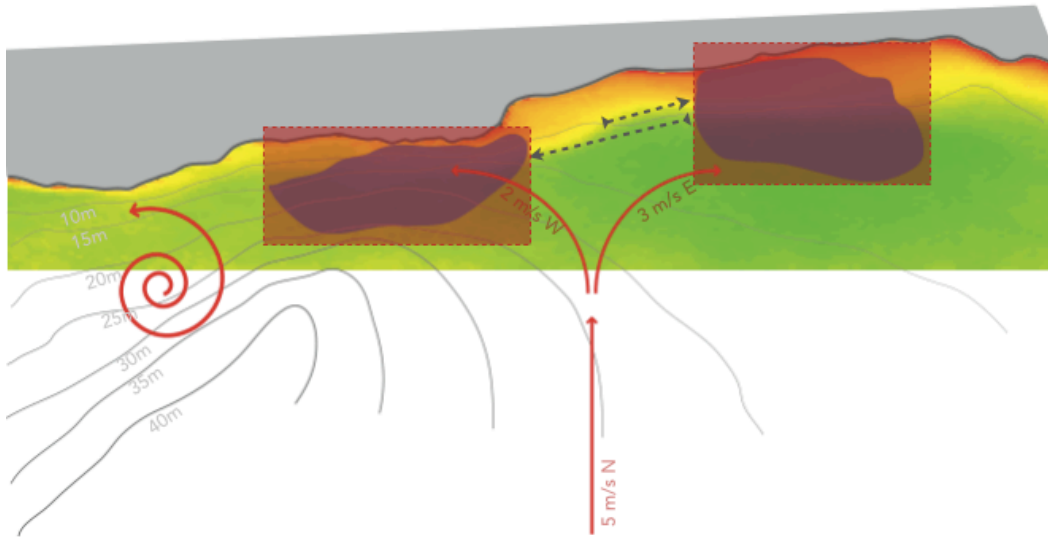
Data visualization

Figure 3: Two putative genetically differentiated populations (blue polygons) overlaid with proposed marine protected area boundaries (red rectangles) and oceanographic data including sea surface temperature, bathymetric contours and major currents and eddies. © Martin Mendez
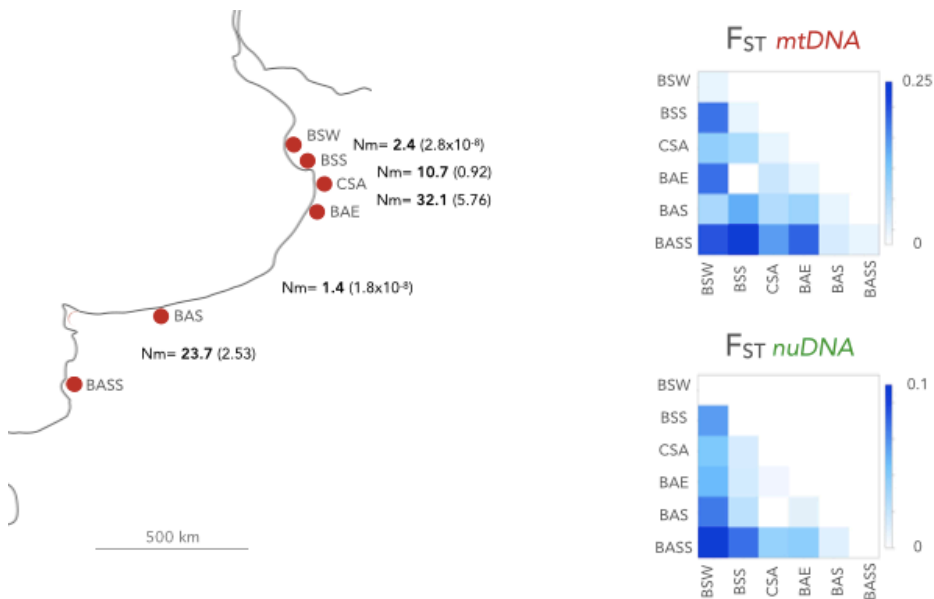

Figure 4: Matrices showing degree of genetic differentiation (far right) for six groups of franciscana dolphin (P. blainvillei) sampled off the coast of Argentina (left) (Mendez et al. 2010) © Martin Mendez
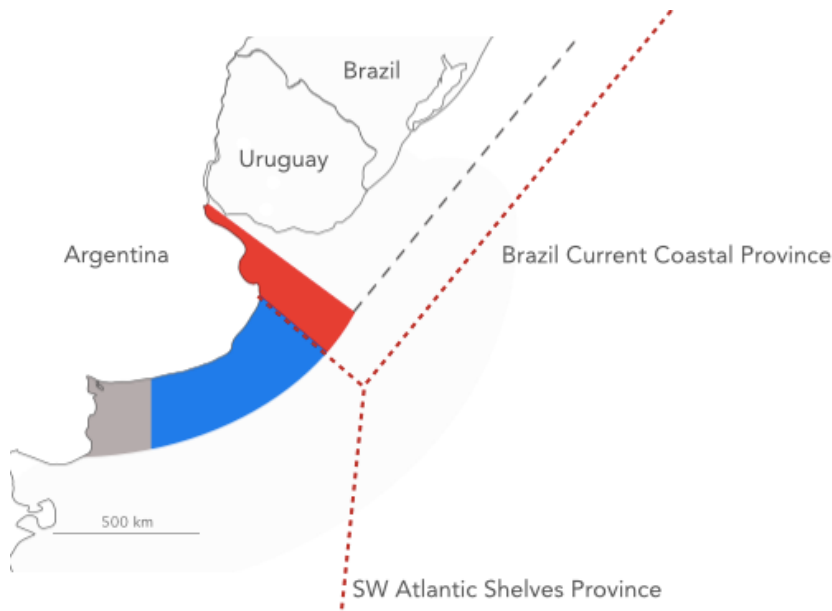
7

Figure 5: The location of three genetically differentiated populations of franciscana dolphin (P. blainvillei) overlaid with the location of oceanographic provinces within Argentina's exclusive economic zone (EEZ). © Martin Mendez

3) Population genetic tools for quantifying connectivity (Oscar Gaggiotti)

Movements of individuals over space provide a web of connectivity among local populations within metapopulations. This leads to complex dynamic dependencies across local population boundaries. Quantifying this connectivity is essential for understanding the spatial dynamics of metapopulations and therefore for proposing their effective management plans for conservation.

There are different types of connectivity: larval connectivity through the dispersion of larvae during their pelagic phase, juvenile/adult connectivity through their daily or seasonal migrations, and overall connectivity that integrates the effect of dispersal at the larval and juvenile/adult stages. Importantly, the type of connectivity genetic methods measure depends on the characteristics of the samples used and the particular statistical methods applied. The traditional population genetics approach to quantify connectivity is based on summary statistics that measure genetic differentiation between populations (i.e., connectivity α 1/genetic differentiation). It is also possible to infer a local measure of genetic isolation using population-specific F_{ST}. Population-specific F_{ST} can be visualized by carrying out spatial interpolation using kriging, and recognizing the reliability of this approach is highly dependent on the sampling scheme employed. Summary statistics are, however, difficult to interpret because they measure the genetic differentiation on an evolutionary timescale, thus they measure *"historic"* connectivity and are biased estimators of *"demographic"* connectivity, making them less relevant for informing marine spatial planning.

An alternative approach for measuring connectivity is the estimation of migration rates. Migration rates are the most biologically meaningful measure of connectivity as it directly affects the spatial dynamics of populations and their response to management. Samples from adults provide an estimate of overall connectivity and samples from larvae provide estimates of larval connectivity. There are two alternative approaches to estimating

8

migration rates: 1) coalescent-based methods that retrace the genealogy of genes using allele frequency or sequence data and provide estimates of migration on an evolutionary timescale (i.e., historic connectivity); and 2) multilocus-genotype methods that assign individuals to local populations and provide estimates of migration on an ecological timescale (i.e., demographic connectivity). Multilocus-genotype methods are based on complex likelihood models that include several parameters and are implemented by computationally intensive Bayesian methods (e.g., BayesAss; BIMr). Using these methods, genetic data can be combined with other sources of information to increase their power or to test hypotheses about the effect of environmental or physical processes. The outputs can be represented by network diagrams, with the nodes representing different populations and the links representing the magnitude and directionality of migration rate estimates.

Data visualization

Figure 6: Spatial interpolation of population specific F_{ST} using kriging for human populations. Sampling locations identified by circles (Foll & Gaggiotti 2006).
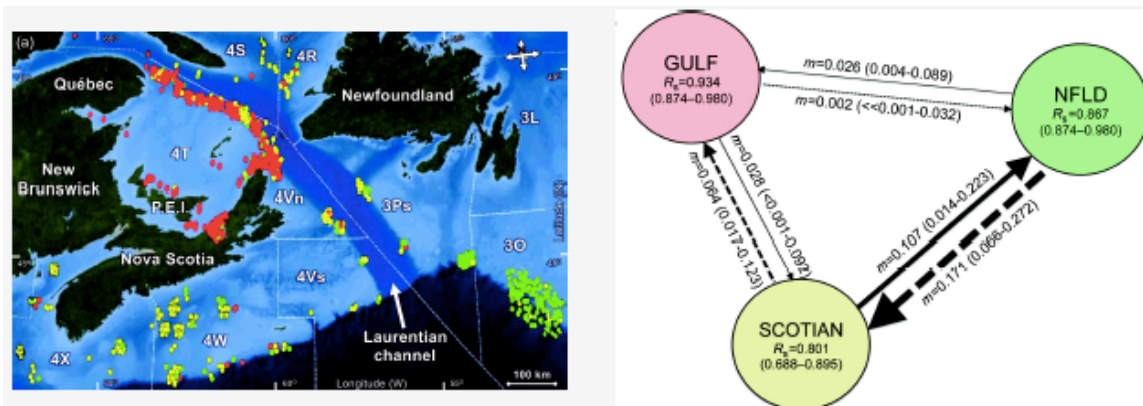

Figure 7: Sampling locations (left) and BayesAss estimations of demographic connectivity between three populations of white hake (right). Magnitude and directionality of the estimated number of migrants (m) are depicted by the size and directionality of the arrows (adapted from Roy et al. 2012).

4) Genetics as a tool for the spatial management of elasmobranchs (Les Noble, Michelle Frost & James Thorburn)

Genetic tools are useful for many different aspects of elasmobranch management and marine protected area planning, including the identification of cryptic species, identification of units/stocks, estimates of gene flow and connectivity, and identification of areas of importance.

The common skate (*Dipturus batis*) is considered a "critically endangered" species and there are international, regional and national protection mechanisms in place for its protection. It was known to be widely distributed but now only exists in remnant populations in NW Scotland, Celtic Sea, and the Rockall Plateau. Historically, the common skate has been considered a single species; however, in 2012 morphometric and molecular analyses revealed the taxon comprised two cryptic species: *D. intermedia* and *D. flossada*. Tagging studies and acoustic monitoring suggest a resident population of common skate in the Firth of Lorn and Sound of Jura region, which led to a proposal for a marine protected area. Genetic analyses of the mitochondrial control region and five nuclear microsatellites demonstrated the presence of a single species in the region (*D. intermedia*), leading to the establishment of a MPA with the common skate as the designation species. Genetic tools are now being used to assess population structure within this stock, and to determine the geographic distribution and spatial ecology of the species complex across the broader region using both genetic and spatial tools to examine genetic diversity, aggregation dynamics and relatedness of individuals.

The spurdog (*Squalus acanthias*) has undergone a 95% reduction in biomass in the NE Atlantic and is now considered critically endangered by the IUCN. Spurdog are also one of the Scottish government's Priority Marine Features (PMFs) and are a priority for future research. Traditionally, spurdog has been considered a highly migratory species, however, fine-scale tagging studies have also suggested a high level of site fidelity. Genetic analyses indicate that the haplotype diversity of this species is relatively evenly represented across samplings sites in different parts of the UK. Genetic analyses provided a useful means to complement local tagging studies by providing broader scale insights. These findings suggest that a local, static MPA can be used in conservation initiatives of this highly mobile species, when the management objective is to maximize the protection of genetic diversity. Genetic tools also provide a means to monitor diversity once an MPA is in place, to assess migration rates between populations within and outside MPAs, determine the movements of individuals through genetic monitoring, and to identify the presence of sink populations.
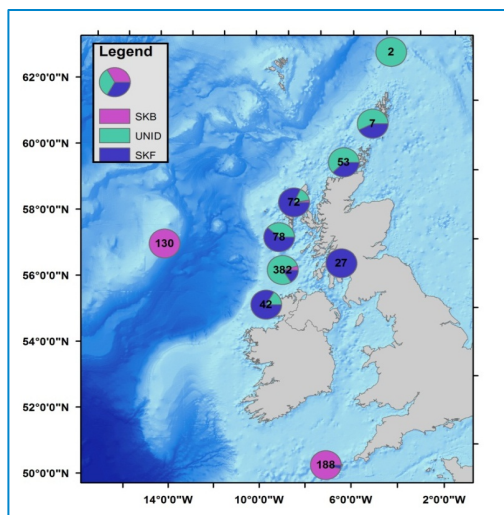
Data visualization:

Figure 8: The number of samples of common skate species: D. intermedia (SKF), D. flossada (SKB), and unidenified (UNID). Samples pooled for each degree of latitude © Michelle Frost & Les Noble
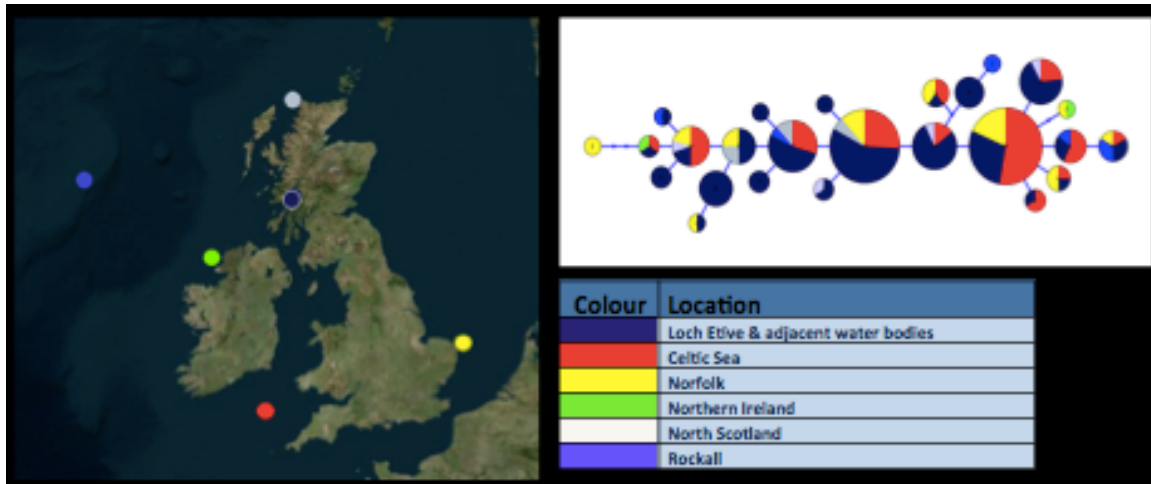
10

Figure 9: Six sampling locations of spurdog (S. acanthias) in the U. K. Corresponding haplotype network (upper right) shows the number of haplotypes (individual circles) and the genetic distance between them (branch length), the number of individuals sharing each haplotype (size of the circle), and the location at which each individual was sampled (colors). © James Thorburn & Les Noble

5) Geodesign using SeaSketch (Will McClintock)

SeaSketch (www.seasketch.org) is a platform for collaborative ocean GeoDesign. Using SeaSketch, anyone with a web browser and internet connection may design management plans, including for marine protected areas and spatial planning. Users can generate hundreds of alternative proposals representing a range of perspectives and interest through an accessible map interface. Zoning, regulatory, or management plans can incorporate the diverse ideas of stakeholders most affected by decisions, in addition to those of the planners and scientists involved. Integrating genetic information into SeaSketch would be an effective first step towards addressing the communication and knowledge gap between geneticists and the marine spatial planning community.

Projects in SeaSketch can be developed worldwide and users are provided with the tools to: 1) initiate a project by defining a study region; 2) upload map layers from existing web services (e.g., human-use, habitat, species distribution, spatially-explicit genetic information); 3) define "sketch classes" such as prospective MPAs; 4) author sketches and receive automated feedback on those designs, such as the ecological value or the potential economic impacts of a MPA; and 5) share sketches and discuss them with other users in a map-based chat forum. SeaSketch also provides analytical feedback about proposed MPAs related to metrics developed to reflect specific management objectives, including the percentage or range of habitats protected, and potential social and economic costs and benefits. Advanced analyses such as Marxan, trade-off methods, and cumulative impacts can also be incorporated. New metrics and analyses for objectives related to genetic measures could potentially be developed and integrated alongside other types of information.
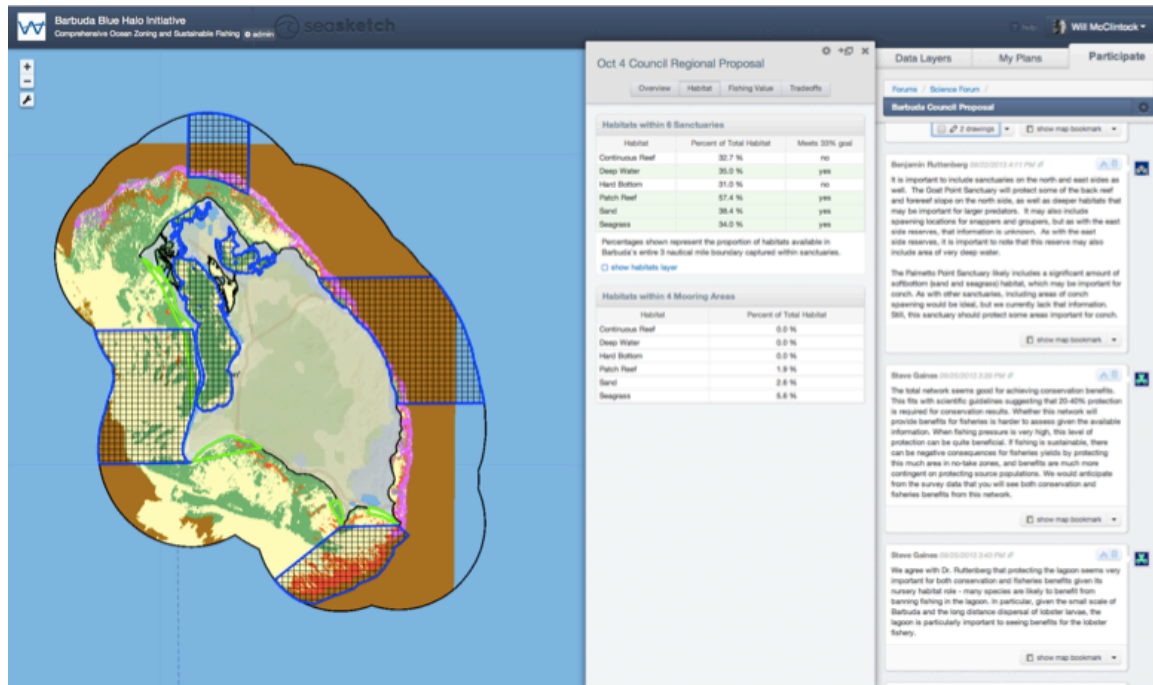
Data visualization:

Figure 10: Example screenshot from SeaSketch showing proposed protected areas (blue crosshatch) overlaid on a habitat map for the island of Barbuda (far left panel), an analytical report of the performance of each proposed area in relation to habitat coverage (central panel), and the online chat forum available for planners (far right). © Will McClintock.

6) GeneGIS – integrating DNA profiles and photo-identification to inform marine spatial planning (C. Scott Baker)

An increasing number of large-scale, collaborative studies of marine and terrestrial megafauna are using spatio-temporally explicit, individual-based records, including photoID catalogues (typically many records) and DNA profiles from biopsy samples (typically few records). The *geneGIS* initiative (www.genegis.org; Dick et al. 2014) in *Wildbook* (www.wildbookdna.org) provides a platform to integrate the location of known individuals in space (via the Structure of Populations, Levels of Abundance, and Status of Humpbacks (SPLASH) and A Pattern of Dolphins (aPOD) databases), a DNA profile for that individual, and locations of environmental covariates in space. Integrating these three types of information can provide insights into population structure and kinship, habitat preference and use, and insights into the ecological processes underlying genetic structure (i.e., seascape genetics).

The *Wildbook* platform reconciles the two sources of identity (photoID and DNA profiles) in a single flat-file format. It also facilitates the spatial display and selection of data, data exploration and analyses, and data export. Data exploration tools include the display of all encounters with humpback whales identified within a defined geographic region, and the sex and DNA profiles of these individuals color-coded by maternal lineage, providing insights into regional differences. Applications of *Wildbook* fall within three primary categories: 1) molecular ecology, enabling comparisons of haplotype frequencies and calculation of genetic differentiation (F_{ST}) between two user-defined regions; 2) capture-

recapture; and 3) kinship analysis, searching all DNA profiles for the closest match to a selected ID enabling the identification of first-order kin. geneGIS also facilitates the conversion of genetic data sets to alternative file formats suitable for a range of analysis programs, thus enabling higher-order genetic analyses.

Data visualization:

Figure 11: Map showing the distribution of humpback whale mtDNA haplotypes across the northern and central Pacific. Five distinct population segments are identified (1-5) © C. Scott Baker
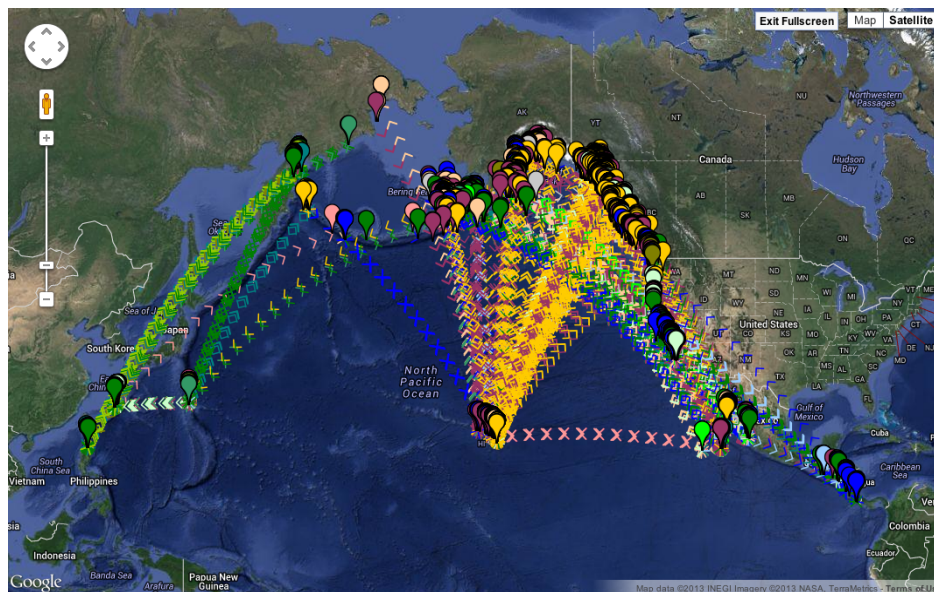

Figure 12: Map showing sampling locations and putative migratory routes of genetically identical individuals (each individual represented by a different color) © C. Scott Baker

13

Brainstorm Session: Why is genetic data/information important for MSP?

- **What are the data/protection gaps that exist if genetic data isn't included in marine protected area and spatial planning?**

- **What types of genetic data are useful for MSP? How is it formatted?**

Genetic tools can provide unique insights pertinent to spatial planning efforts (e.g., marine protected area design) that have a variety of goals, e.g., single-species management, fisheries management, critical habitat protection, or ecosystem management. The integration of genetic data and information into the management process needs to be dynamic and adaptive based on changing management objectives. These objectives may subsequently require updating based on the genetic information that is being considered (Figure 13).

Figure 13: The dynamic relationship between the representation and integration of genetic information into the management process.

This information falls into three primary categories: 1) Defining units; 2) Assessing connectivity; and 3) Measuring status and trends:

*See Appendix 1 for a summary of this (and the next) section.

1) Defining Units:

Genetic tools can provide a means of quantifying similarity and differentiation between units at multiple scales, i.e. individuals and kin-relationships, demographic populations, species, and communities/ecosystems.

- Individuals and kin-relationships
    - ⇒ Type of data: microsatellites
    - ⇒ Data format: probability of identity, probability of parentage/kinship, relatedness measures, pedigrees, relatedness networks

- Demographic populations
    - ⇒ Type of data: microsatellites, single nucleotide polymorphisms, mitochondrial sequences
    - ⇒ Data format: genetic differentiation indices (e.g., F_{ST}, ϕ_{ST}), clustering algorithms (e.g. STRUCTURE), assignment probabilities (e.g., GENELAND), PCA-based analyses

- Species
    - ⇒ Type of data: nuclear gene sequences, nuclear introns, mitochondrial sequences, genome
    - ⇒ Data format: phylogenetic trees
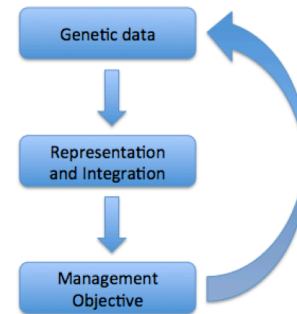
- Community/Ecosystem

⇒ Type of data: nuclear gene sequences, nuclear introns, mitochondrial sequences, genome, environmental DNA
⇒ Data format: phylogenetic trees, clustering algorithms, species and community diversity/richness/evenness indices

2) Assessing connectivity

Related to the definition of units, genetic tools can be used to assess the connectivity between units defined *a priori* using the methods described above. Migration rates are one of the most biologically meaningful measures of connectivity as they directly affect the spatial dynamics of populations, partly due to the fact that they can be influenced by environmental factors. There are two alternative approaches to estimating migration rates using genetic tools: 1) coalescent-based methods that retrace the genealogy of genes and provide estimates of migration on an evolutionary timescale (i.e., historic connectivity); and 2) multilocus-genotype methods that assign individuals to local populations and provide estimates of migration on an ecological timescale (i.e., demographic connectivity). Direct movements of individuals can also be measured using genetic capture-recapture methods, providing a real-time measure of connectivity.

- Historic connectivity
  ⇒ Type of data: mitochondrial control region sequences, microsatellites
  ⇒ Data format: network, connectivity matrix

- Demographic connectivity
  ⇒ Types of data: microsatellites
  ⇒ Data format: network, connectivity matrix

- Individual Movements
  ⇒ Types of data: microsatellites
  ⇒ Data format: genotypic matches (i.e., genetic capture-recapture)

3) Measuring Status and Trends

Genetic tools also offer a means of measuring the status of the units identified using the methods above, and to assess how the status of these units changes through time. There are a number of measures that would be useful in this regard, including the ability to measure population reduction including the detection of historic population bottlenecks, effective population size, population diversity, and species diversity. It is important to note that status indicators are useful only as relative measures when comparing a suite of populations of the same species. For comparison of data for different species (i.e., comparing populations of different cetacean species in a given area), some information on demographic parameters (e.g., generational time, age at reproductive maturity, etc.) would also be needed.

To understand trends, it is important to consider the temporal scale and baseline that is most useful for addressing the management question, as this will determine the sampling scheme required. Gaining a sense of historical trends can be captured with a one-time sampling approach, while understanding current trends by comparing measures of Ne across a time period would require the periodical collection of samples (e.g., every 5 years).

- Population reduction and bottlenecks
    ⇒ Type of data: microsatellites, DNA sequences
    ⇒ Data format: allele frequency distribution, time since bottleneck, change in population size over time

- Effective population size (Ne) (linked to mating systems)
    ⇒ Type of data: DNA sequences, microsatellites
    ⇒ Data format: Ne (continuous metric)

- Population diversity
    ⇒ Type of data: microsatellites, single nucleotide polymorphisms, mitochondrial sequences
    ⇒ Data format: diversity indices (continuous metric): e.g., allelic richness, nucleotide diversity, etc.

- Species diversity
    ⇒ Type of data: nuclear gene sequences, nuclear introns, mitochondrial sequences, genome, environmental DNA
    ⇒ Data format: phylogenetic trees, clustering algorithms, species and community diversity/richness/evenness indices

Breakout Session: How can genetic data/information be integrated into MSP tools?

1. **How can genetic data/information be visualized within a MSP tool in a way that improves its <u>ecological meaning</u> and <u>usefulness</u> for the marine conservation and management community?**

2. **What are the challenges and potential pitfalls?**

*See Appendix 1 for a summary of this (and the previous) section.

Group 1: Units

When determining how to visualize units in a marine spatial planning tool, two types of units may need to be considered: 1) *a priori* units (including individuals); and 2) Non *a priori* units.

1) *a priori* units

Species/Community level: At the species level, the number of species in a defined area is often a key criterion of a marine management plan. The representation of species level genetic diversity is therefore an important measure to include in marine spatial planning. For example:

    ⇒ Levels of species diversity within a defined area (e.g., Figure 14)
    ⇒ Diversity hot-spots or cold-spots

Population level: At the population level, the most common way to identify whether pre-defined units are demographically distinct or not is to employ a suite of differentiation indices, such as population-specific F_{ST}, pairwise F_{ST}, ϕ_{ST}, Jost's D, etc (e.g., Figure 13).

These indices of genetic differentiation between *a priori* units can be visualized in a marine spatial planning tool using approaches such as:

⇒ Levels of genetic diversity for different populations within a species
⇒ Kernel distribution of genetic differences
⇒ Haplotype distributions
⇒ Heat map / kriging of population-specific F_{ST}
⇒ Sliding window analysis
⇒ Principal component analysis
⇒ Isolation by distance
⇒ Ecological analysis (e.g., isolation by environmental distance, Mantel tests)

However, since the reliability of some interpolation methods previously cited is highly dependent on the sampling scheme employed (e.g. .the local distribution of samples) and the local environmental conditions (e.g. the presence of hydrographic front or others heterogeneous environmental structures), an associated output error map can be created to visualize the degree of uncertainties of the estimates (e.g. the standard deviation or confidence interval maps).
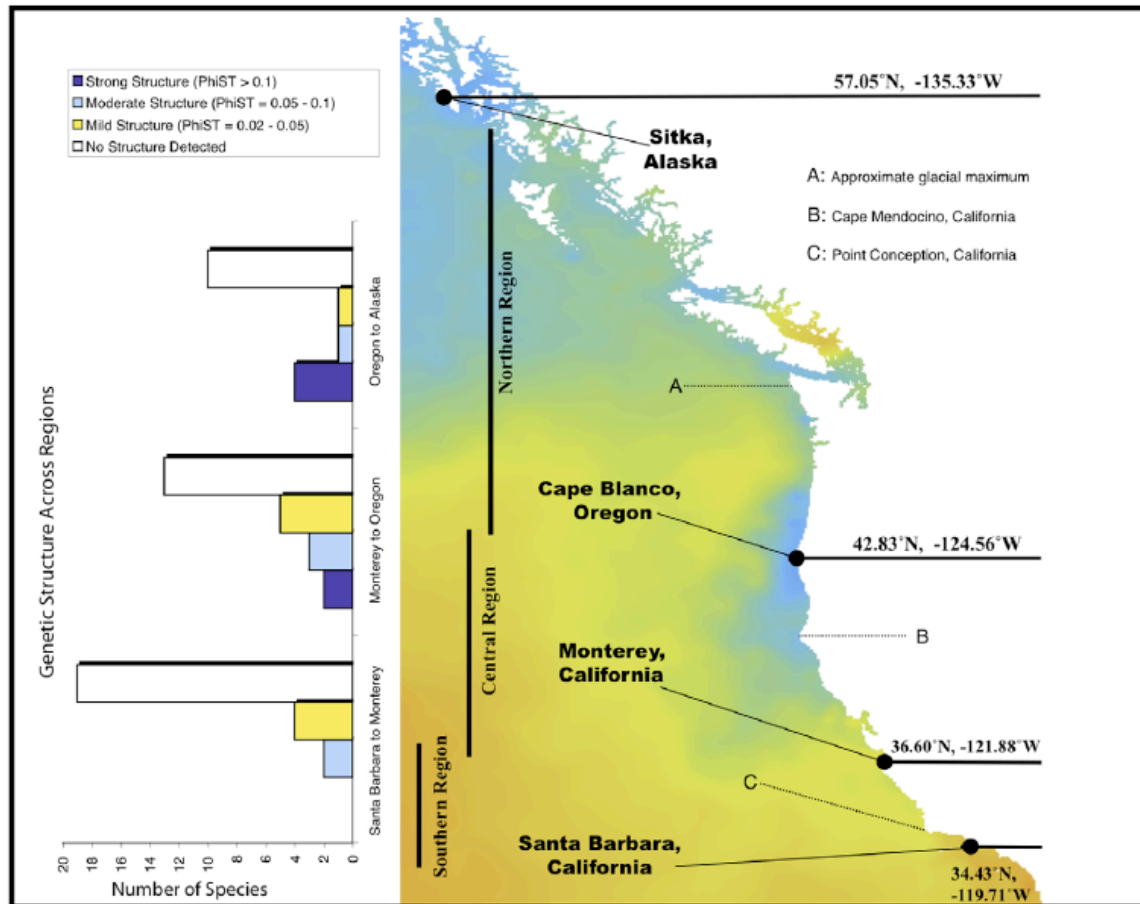
Figure 14: Visual representation of genetic structure of near-shore invertebrate species in the Pacific coast of North America. Bar charts on the left-hand side of the figure show the number of species with strong, moderate and mild genetic structure in southern, central and northern comparisons. Structure categories are mild: ΦST = 0.02 – 0.05; moderate: φST = 0.05 – 0.10; strong: φST >0.10. Background color represents sea surface temperature for a day in July, 2008 (Kelly & Palumbi 2010).

Individual: Individuals are, by *de facto*, identified *a priori*. The range size of individuals and their relatedness/kinship within defined areas, for example, may provide useful insights into the delineation of units of management. Approaches to representing this information in a marine spatial planning tool can include:

⇒ Individual kernel distributions (interpreting genetic data alongside satellite telemetry)
⇒ Relatedness/kinship measures (continuous metric)

2) Non *a priori* units

Species/Community level: When the boundaries of species are not clear (e.g., when hybridization may be in occurrence or when species are cryptic/not easily identifiable when sampling), phylogenetic analysis can be undertaken to identify different species or other evolutionary significant units. These units can then be visualized in the same manner as described for *a priori* units.

Environmental DNA analyses can also be employed for a rapid assessment of diversity at the community level in some systems (e.g. the sea bed). In a marine spatial planning tool, this information can be visualized as:

⇒ Similarity/differentiation indices between pre-defined areas (continuous metric)
⇒ Clustering of units sampled in different areas represented by spatial interpolation

Population level: Clustering methods can be used to detect genetically different groups within a sample when populations cannot be identified *a priori* or to act as an independent test of those predefined groups:

⇒ Clustering analysis of individuals sampled in different areas represented by spatial interpolation

Group 2: Connectivity

For the purposes of management and marine spatial planning, demographic connectivity was considered the most relevant, as it provides estimates of recent connectivity between management units and also retention of individuals within a single unit. Movements of individuals identified by genetic capture-recapture analysis (the identification of the same individual in different locations and/or different years using genetic methods) can be used as an indicator of real-time movements and therefore is also of relevance to management.

1) Demographic connectivity

Demographic connectivity can be estimated using methods that assign individuals to local populations using multilocus genotype data and provide estimates of migration. Multilocus-genotype methods are based on complex likelihood models that include several parameters and are implemented by computationally intensive Bayesian methods (e.g., BayesAss; BIMr). Using these methods genetic data can be combined with other sources of information to increase their power or to test hypotheses about the effect of environmental or physical processes. The outputs of these methods can be represented either as network diagrams or connectivity matrices.

Network diagrams: Depict different management units and the estimates of migration between them as a network diagram. The nodes of the network represent each different management unit and the links connecting them represent the magnitude and directionality of the migration rate estimates. If there are few management units under consideration, these networks can either be mapped spatially onto other data layers, or as an accompanying analytic (see Figure 7). Visualization of connectivity as a network is challenging when there are many management units under consideration, however; if this is the case, a connectivity matrix may be a more appropriate (see below).

As marine protected area and spatial planning is primarily concerned with the degree of connectivity between units, rather than the mapping of movement 'corridors', the representation of connectivity as straight lines between units was deemed adequate (as opposed to, for example, raster based layers depicting more complex spatial analysis such as least-cost paths).

Connectivity matrix: When the number of management units is high, estimates of connectivity can be represented as a pairwise matrix, with the rows and columns representing the different management units, and the data representing the migration rate estimates (see Figure 15). Such a pairwise matrix is useful as it can also be used to display estimates of retention within the same management unit along the diagonal. Values can also be color-coded to aid interpretation.

In either the case of network diagrams or connectivity matrices, it would be useful to easily identify when connectivity estimates fall below a certain threshold value (e.g., less than one mating individual every other generation) that may be stated in the management plan/objectives. This tool is commonly used in ecological monitoring, and the information is gathered in "biodiversity indicators dashboard", giving useful information on biodiversity trends, with the objective to be highly comprehensive for policy-makers and managers. This can easily be addressed by incorporating a toggle or slide tool into the marine spatial planning tool, which would enable users to easily observe how levels of connectivity change with different threshold values.

	To			
	Pop1	**Pop2**	**Pop3**	**Pop4**
Pop1	0.9	0.4	0.7	0
Pop2	0	1	0	0.5
Pop3	0.5	0.2	0.7	0.3
Pop4	0.5	0.2	0.3	0.8

Figure 15: Example of a connectivity matrix for four management units. Values represent migration estimates to and from each management unit. Retention is displayed along the main diagonal.

Genetic capture-recapture: Genetic profiling allows the identification of the same individual sampled at multiple times and locations. This can provide useful insights into direct (i.e. real-time) connectivity between locations and patterns of site fidelity. Genetic profiling also enables the exploration of these patterns among kin-groups.

Genetic capture-recapture information can be visualized in a number of ways:

⇒ Straight lines connecting two different sampling locations (may be misleading if the line is interpreted as being indicative of the path of movement the individual undertook between locations).
⇒ Individual range maps, kernel distributions, or densities (interpreting genetic data alongside satellite telemetry).
⇒ Polygons proportional to the number of times the individual was recaptured. If the same individual is recaptured in the same location at different times, a simple polygon (e.g., circle) could be used to represent the individual at the location and sized proportionally to the number of times the individual was recaptured (e.g. 2x, 5x, 10x). The same technique can also be used to visualize juvenile - parentage assignments (Figure 16).
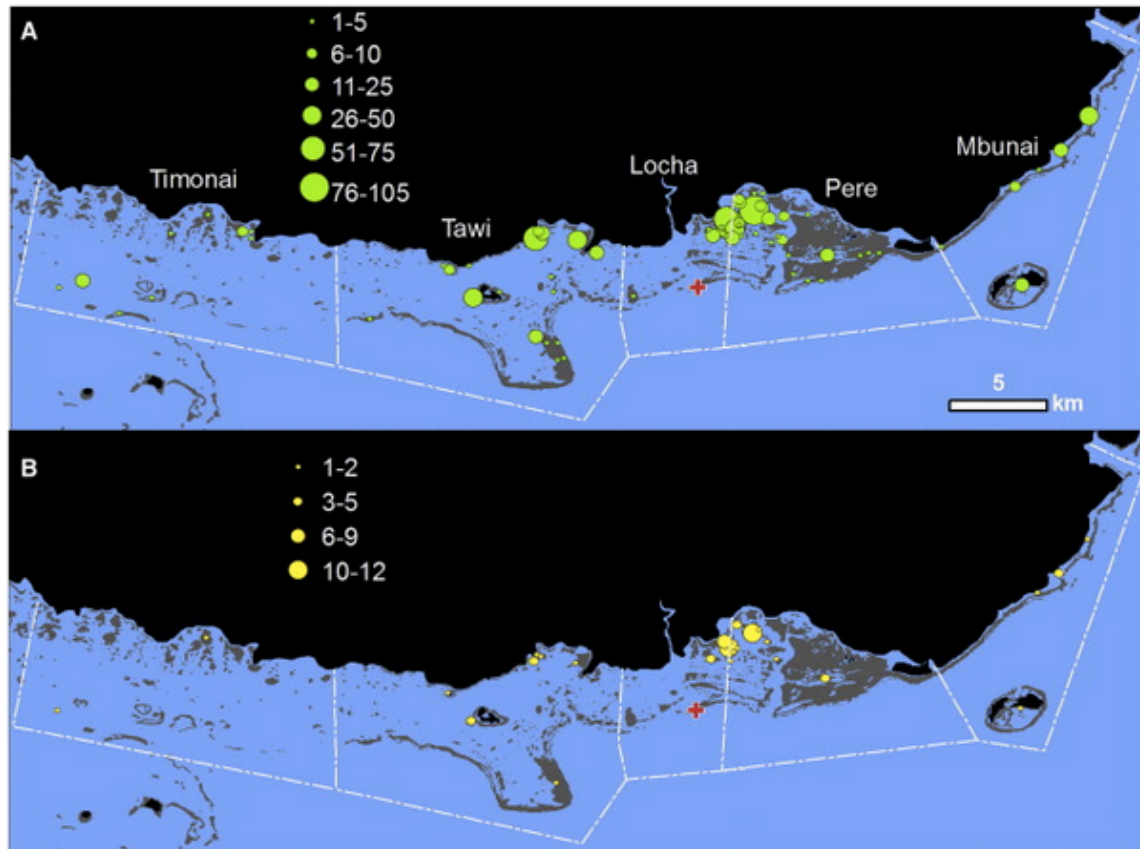
20

Figure 16: Spatial patterns of coral grouper (Plectropomus areolatus) in Papua New Guinea. (A) juvenile sample collection and (B) juvenile parentage assignments. Green (A) and yellow (B) circles are scaled to the number of juveniles. Adults were sampled from a single fish spawning aggregation (red cross), and juveniles were collected from 66 individual reefs (green circles in A). White dashed lines are customary marine tenure boundaries of the five communities. Coral reefs are depicted in gray (Almany et al. 2013).

Group 3: Genetic diversity and effective population size (Ne)

Genetic diversity refers to the total number of genetic characteristics in the genetic makeup of the individual. The genetic diversity of a population directly related to demographic processes, such as population reduction or expansion (see section on Ne, below). Higher relative levels of genetic diversity therefore become a useful indicator that populations will be more resilient to surviving potential bottlenecks (i.e., rapid reductions in population size) if their diversity is higher, as they have a lower probability of permanently losing alleles. It follows that populations with lower relative levels of genetic diversity may be more vulnerable to extinction if the size of the population is reduced.

Genetic diversity is the material upon which evolution operates and so genetically diverse populations are also hypothesized to be generally more resilient and able to adapt to environmental change. Evidence from epigenetic screening and outlier loci under selection indicate that this relationship does not always hold, however, and more research is needed to explore the differences between neutral genetic markers and those under selection, and also at the epigenetic level. Increasing access to next generation sequencing (NGS)

technologies will enable conservation geneticists to place more emphasis on identifying selectable genes/markers in genomic data as an indicator of population health.

In terms of marine spatial planning, the quantification of genetic diversity for the management units identified may therefore provide a useful way to prioritize management action depending on the goals of the management plan (e.g., higher protection for management units with low genetic diversity).

Levels of genetic diversity are inextricably linked to effective population size (Ne), or the number of breeding individuals in an idealized population that would show the same amount of dispersion of allele frequencies under random genetic drift or the same amount of inbreeding as the population under consideration. More simply, the term generally refers to the number of breeding individuals within a population per generation. Populations with large census size (i.e., total number of individuals) can have small effective population sizes, and vice versa, meaning that estimates of Ne can often not be intuited from the census size alone (although Ne is often considered to be 10-20% of the census size (Luikart et al. 2010)). Coalescence-based modeling of Ne can provide insights into how the Ne of the population has changed through time, for example, if a historically large population has recently undergone a rapid reduction in size (i.e., a genetic bottleneck). This again provides useful information for management prioritization decisions. Next generation sequencing (NGS) data and techniques that employ linkage-tract lengths (e.g., Gusev et al. 2012) will likely improve the performance of such approaches.

Most measures of genetic diversity and Ne can be described by a simple summary statistic. Therefore, in terms of visualization, the value of the statistic could be used to color-code a polygon that represents the management unit, once its boundaries have been identified. There are a number of options that could be considered when deciding on this approach:

⇒ Raw variables: each measure of genetic diversity (e.g., effective number of alleles, number of haplotypes, haplotype diversity, nucleotide diversity, etc.) could be attributed a separate polygon and color-coded. The value of the measure could then correspond to the intensity of the shading of the polygon.
⇒ Standardized variables: each measure of diversity could be converted to a standardized scale (i.e., 0-1) prior to being color-coded. This might make the values easier to interpret by the user.
⇒ Composite variable: each measure could be converted to a standardized scale and then integrated into a single composite measure of "genetic diversity". This would be simplest in terms of interpretation by the user, but may be challenging to derive in a meaningful way as not all component measures are directly correlated (i.e., two populations with the same composite value may have different component measures).

There are a number of different ways to calculate metrics of diversity and effective population size, and so it will be important to adopt a standard approach for all data included in the marine spatial planning tool (or it will be necessary to incorporate a conversion factor following raw data entry).

Irrespective of the approach, it is very important for the user to also have access to the levels of uncertainty corresponding to each metric. While there may be ways to visualize this, it was concluded that the visualization of the raw or standardized metric was the most

important and that uncertainty information could be supplied as the metadata/attribute information related to each polygon.

Challenges to consider

During the Focus Group, a number of general challenges were identified related to integrating and visualizing genetic information in a marine spatial planning tool. Each challenge is described briefly below, but each will need to be considered in depth as the approach continues to be developed.

1) Spatial and temporal sampling

The results generated and conclusions drawn from any genetic analyses are highly sensitive to the sampling scheme adopted. How individuals were sampled spatially and temporally directly influences the determination of population/management units, inferences regarding the dynamism of those units through time, and estimates of genetic diversity. These issues are particularly pertinent to studies of highly mobile and migratory species, where a genetic study may only capture a snapshot of the patterns existing at that particular stage of their migration or life cycle.

In light of the fact that comprehensive spatial and temporal sampling is not possible, it is very important that details of the sampling scheme and associated levels of uncertainty be made available to the user of the marine spatial planning tool in a way that is very clear to understand. If not, there is a very high risk that management decisions will be made based on highly uncertain or inappropriate data that could be of serious detriment to the management units under consideration. As most users of marine spatial planning tools are not geneticists, significant effort is needed to explore how this issue can be addressed so that subsequent management decisions are biologically and ecologically meaningful.

2) Extent and resolution

Related to the issue of sampling, is that of extent and resolution of the data used. Inferences regarding the delineation of management units and their connectivity may be very different when a study region is 10 km^2 compared to 100 km^2. Moreover, this issue is also introduced by the sampling scheme employed. Spatial resolution will be determined between individuals or groups of individuals, and the number of individuals per group. The genetic marker used (e.g., sequences, multi-locus data) for these samples will also determine the resolution of the analysis. A robust methodology to scale our understanding of management units both up and down has not yet been developed and should be a focus of significant future research. In the interim, explaining this issue clearly to the marine spatial planner and/or incorporating this into a measure or metric of uncertainty should be undertaken.

3) Thresholds of "difference"

Management plans often require information regarding thresholds of "difference", for example, are management units different or not, are they adequately connected or not, are they genetically diverse or not? As forms of genetic data and information are generally continuous, there is a need to incorporate the ability to explore different thresholds into the marine spatial planning tool, so that users can make inferences based on thresholds defined in their specific management plan, or to enable the exploration of how genetic inferences

change across a range of thresholds. The simplest way to incorporate this into a marine spatial planning tool would be to employ levels of statistical significance, such as those used in analyses of hierarchical variance (e.g., analysis of population structure using F-statistics) to indicate if sampling units are genetically different. In cases where statistical significance cannot be derived, it would be useful to develop a "sliding scale" tool for each relevant genetic metric, which would allow the user to define threshold values *a priori*, or explore how genetic patterns vary with different thresholds. Developing methods to aggregate or combine different thresholds for different taxa in a meaningful way will require more research. It was concluded during the Focus Group that these thresholds should be defined as part of the management plan, as they are specific to species and management goals, rather than defined by developers of the marine spatial planning tool itself.

4) Uncertainty

Every aspect of genetic analyses carries some component of uncertainty and it is essential that users of marine spatial planning tools have this uncertainty clearly presented to them so that it can be considered in parallel with the results of the genetic analyses. It was concluded during the focus group that uncertainty information could be supplied as the metadata/attribute information. However, the challenge lies in communicating this information adequately and clearly to marine spatial planners, so that they both understand the information and include it in any management decisions that are made. An immediate way to illustrate these uncertainties is through the use of synthetic map that can be considered as supplementary information in the decision process. A specific consultation process with managers is recommended to develop an approach to addressing this.

5) Combining the whole information into a simple and comprehensible index

As this report demonstrates, there is ample genetic data and information that can be spatially evaluated. However, marine spatial planning needs to consider all these data simultaneously rather than individually. Combining all this information in an intelligible way is a necessity to help managers to make decisions. Integrating the total available genetic information may be achieved by creating an index based on the various information calculated for each area. These index values could be easily mapped to highlight areas of major importance and provide a useful support tool for managers charged with identifying EBSAs.

6) Predictive framework

Current genetic analyses provide important information for management on species and population patterns, however methods provide a current or historic snapshot rather than enabling predictions of how these patterns may change in the future. Given the need to adaptively manage in the face of environmental change, the ability propose and test scenarios and to make predictions would be highly useful for managers. There is therefore a significant need to advance scientific understanding of the mechanisms (i.e., biological, ecological, and evolutionary processes) that are driving the genetic patterns observed. Understanding these mechanisms represents an important initial step to developing a predictive evolutionary framework. Ecological based models that characterize important habitats and prey availability will likely be useful in this context.

Conclusions

Genetic data can inform marine spatial planning in a variety of ways at the species, population and individual level that are not captured by other types of data and information. There is therefore a significant need to ensure that this data is considered in a systematic way during the marine spatial planning process. Interpretation of genetic analyses can be challenging for those unfamiliar with the field, however, and so it is essential that genetic data can be visualized in a way that is understandable and relevant for marine spatial planners.

The outcomes of the initial effort to develop a standardized approach to integrating genetics into marine spatial planning that took place during this Focus Group demonstrate that there are a wide variety of ways that genetic information can be visualized and integrated successfully into a marine spatial planning tool at multiple scales. It is hoped that the ideas put forth in this report form the basis of further discussion, testing, and refinement of these methods, which will result in a standardized approach that can be widely adopted.

The discussions also highlighted a number of challenges in the interpretation of the genetic information once it is visualized (e.g. uncertainty, sampling bias, threshold values) that need to be thoroughly addressed as part of the development of this approach. Educating marine spatial planners on the uncertainty of the data, whilst also providing useful information upon which to base decisions represents a key challenge that we recommend is specifically addressed through consultation with planners.

Finally, an important conclusion from the discussions was that not all types of genetic information can be spatially visualized, however they are still extremely valuable and should be integrated into management decisions. This represents a challenge for both geneticists and planners and will require the development of close collaboration between the two communities. The development of applied regional or species-specific case studies represents a useful next step towards the development of these collaborations, and the testing and refinement of the methods and challenges summarized in this document.

References

Almany GR, Hamilton RJ, Bode M, et al. (2013). Dispersal of grouper larvae drives local resource sharing in a coral reef fishery. *Current Biology*, **23**:626-630.

Beger M, Selkoe KA, Treml E, et al. (2014). Evolving coral reef conservation with genetic information. *Bulletin of Marine Science*, **90**:159-185.

Dick DM, Walbridge S, Wright DJ, et al. (2014). geneGIS: Geoanalytical tools and Arc marine customization for individual-based records. *Transactions in GIS*, **18**:324-350.

Foll M & Gaggiotti O (2006). Identifying the environmental factors that determine the genetic structure of populations. *Genetics*, **174**:875-891.

Gusev A, Palamara PF, Aponte G, et al. (2012). The architecture of long-range haplotypes shared within and across populations. *Molecular Biology and Evolution*, **29**:472-486.

Kelly RP & Palumbi SR (2010). Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS ONE*, **5**:e8594.

Luikart G, Ryman N, Tallmon D, et al. (2010). Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics*, **11**:355-373.

Mendez M, Rosenbaum HC, Wells RS, et al. (2010). Isolation by environmental distance in mobile marine species: molecular ecology of franciscana dolphins at their southern range. *Molecular Ecology*, **19**:2212-2228.

Roy D, Hurlbut TR & Ruzzante DE (2012). Biocomplexity in a demersal exploited fish, white hake (*Urophycis tenuis*): depth-related structure and inadequacy of current management approaches. *Canadian Journal of Fisheries and Aquatic Sciences*, **69**: 415-429.