Merging Approaches to Explore Connectivity in the Anemonefish, *Amphiprion bicinctus*, along the Saudi Arabian Coast of the Red Sea

Dissertation by

Gerrit B. Nanninga

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

King Abdullah University of Science and Technology

Thuwal, Kingdom of Saudi Arabia

*September 2013*
EXAMINATION COMMITTEE APPROVALS FORM

The dissertation/thesis of Gerrit Nanninga is approved by the examination committee.

Committee Chairperson: Dr. Michael Berumen

Committee Members:  Dr. Ulich Stingl
  Dr. Ibrahim Hoteit
  Dr. Simon Thorrold
ABSTRACT

Merging Approaches to Explore Connectivity in the Anemonefish, *Amphiprion bicinctus*, along the Saudi Arabian Coast of the Red Sea

Gerrit Nanninga

The field of marine population connectivity is receiving growing attention from ecologists worldwide. The degree to which metapopulations are connected via larval dispersal has vital ramifications for demographic and evolutionary dynamics and largely determines the way we manage threatened coastal ecosystems. Here we addressed different questions relating to connectivity by integrating direct and indirect genetic approaches over different spatial and ecological scales in a coral reef fish in the Red Sea. We developed 35 novel microsatellite loci for our study organism the two-band anemonefish *Amphiprion bicinctus* (Rüppel 1830), which served as the basis of the following approaches. First, we collected nearly one thousand samples of *A. bicinctus* from 19 locations across 1500 km along the Saudi Arabian coast to infer population genetic structure. Genetic variability along the northern and central coast was weak, but showed a significant break at approximately 20°N. Implementing a model of isolation by environment with chlorophyll-a concentrations and geographic distance as predictors we were able to explain over 90% of the genetic variability in the data ($R^2 = 0.92$). For the second approach we sampled 311 (c. 99%) putative parents and 172 juveniles at an isolated reef, Quita al Girsh (QG), to estimate self-recruitment using genetic parentage analysis. Additionally we collected 176 juveniles at surrounding locations to estimate larval dispersal from QG and ran a biophysical dispersal model of the system with real-
time climatological forcing. In concordance with model predictions, we found a complete lack (c. 0.5%) of self-recruitment over two sampling periods within our study system, thus presenting the first empirical evidence for a largely open reef fish population. Lastly, to conceptualize different hypotheses regarding the underlying processes and mechanisms of self-recruitment versus long-distance dispersal in marine organisms with pelagic larval stages, I introduce and discuss the concept of “origin effects”, providing the theoretical background to some of the questions that have arisen during this research. Overall, this thesis has generated significant new insights into the patterns of coral reef fish connectivity, specifically for the Red Sea, where such information has previously been scarce.
ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Berumen, and my committee members, Dr. Stingl, Dr. Hoteit and Dr. Thorrold. Specific thanks go to Dr. Saenz-Agudelo and Dr. Manica for guidance and support throughout the course of this research.

I would like to thank all of the existing and former Reef Ecology Lab as well as the Coastal and Marine Resources Core Lab for assistance during fieldwork operations.

Finally, I would like to thank my wife and my mom!
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXAMINATION COMMITTEE APPROVALS FORM</td>
<td>2</td>
</tr>
<tr>
<td>COPYRIGHT PAGE</td>
<td>3</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>4</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>6</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>7</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>8</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>9</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>11</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>12</td>
</tr>
<tr>
<td>CHAPTER 2: REVIEW</td>
<td>43</td>
</tr>
<tr>
<td>CHAPTER 3: PRIMER DEVELOPMENT</td>
<td>74</td>
</tr>
<tr>
<td>CHAPTER 4: POPULATION GENETICS</td>
<td>84</td>
</tr>
<tr>
<td>CHAPTER 5: PARENTAGE ANALYSIS</td>
<td>115</td>
</tr>
<tr>
<td>CHAPTER 6: ORIGIN EFFECTS</td>
<td>137</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>176</td>
</tr>
<tr>
<td>APPENDIX 1: LIST OF PUBLICATIONS</td>
<td>182</td>
</tr>
<tr>
<td>APPENDIX 2: NON-THESIS RELATED FIELDWORK</td>
<td>185</td>
</tr>
<tr>
<td>APPENDIX 3: LIST OF CONFERENCES</td>
<td>186</td>
</tr>
</tbody>
</table>
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>chl-a</td>
<td>Chlorophyll a</td>
</tr>
<tr>
<td>CMS</td>
<td>Connectivity Modeling System</td>
</tr>
<tr>
<td>FDR</td>
<td>False Discovery Rates</td>
</tr>
<tr>
<td>GLM</td>
<td>Generalized Linear Model</td>
</tr>
<tr>
<td>HE</td>
<td>Expected Heterozygosity</td>
</tr>
<tr>
<td>HO</td>
<td>Observed Heterozygosity</td>
</tr>
<tr>
<td>HWE</td>
<td>Hardy-Weinberg Equilibrium</td>
</tr>
<tr>
<td>IBD</td>
<td>Isolation by Distance</td>
</tr>
<tr>
<td>IBE</td>
<td>Isolation by Environment</td>
</tr>
<tr>
<td>IBM</td>
<td>Individual based Model</td>
</tr>
<tr>
<td>LD</td>
<td>Linkage Disequilibrium</td>
</tr>
<tr>
<td>MITgem</td>
<td>MIT general circulation model</td>
</tr>
<tr>
<td>MMRR</td>
<td>Multiple Matrix Regression with Randomization</td>
</tr>
<tr>
<td>MPA</td>
<td>Marine protected Area</td>
</tr>
<tr>
<td>MST</td>
<td>Minimum Spanning Tree</td>
</tr>
<tr>
<td>PLD</td>
<td>Pelagic Larval Duration</td>
</tr>
<tr>
<td>QG</td>
<td>Quita al Girsh</td>
</tr>
<tr>
<td>SR</td>
<td>Self-Recruitment</td>
</tr>
<tr>
<td>uH</td>
<td>unbiased Heterozygosity</td>
</tr>
</tbody>
</table>
LIST OF ILLUSTRATIONS

Chapter 1
Figure 1: Image of *Amphiprion bicinctus* ................................................................. 28

Chapter 3
Figure 1: Sampling locations ....................................................................................... 77
Figure 2: Heterozygosity and pairwise relatedness ...................................................... 82

Chapter 4
Figure 1: Sampling sites .............................................................................................. 89
Figure 2: Unbiased heterozygosities .......................................................................... 97
Figure 3: MST and STRUCTURE analyses ................................................................. 98
Figure 4: Functions of IBD ......................................................................................... 99
Figure 5: MMRR plots ................................................................................................. 100

Appendix A: Red Sea bathymetry and chl-a ............................................................. 112

Appendix C: ln P(K) ................................................................................................... 114

Chapter 5
Figure 1: Sampling sites and assignments ................................................................. 119
Figure 2: Within-cohort heterozygosity and pairwise relatedness ................................ 124
Figure 3: Model predictions ...................................................................................... 125

Appendix A: Anemones at Quita al Girsh ................................................................. 135
Appendix B: Number of alleles and heterozygosities across loci ............................... 136
Chapter 6

Figure 1: Flow chart of origin effects................................................................. 155

Figure 2: Self-recruitment vs. local retention.................................................. 158

Figure 1: Diagram of dispersotypes................................................................. 161
LIST OF TABLES

Chapter 3

Table 1: Microsatellite information................................................................. 78

Chapter 4

Table 1: Sampling regime............................................................................... 91
Table 2: Additional microsatellites................................................................. 92
Table 3: AMOVA results................................................................................ 96

Appendix B: Alternative measures of genetic distance................................. 113

Chapter 5

Table 1: Summary statistics of similar studies.............................................. 117
Introduction:

Coral reef ecosystems are discontinuous shallow water environments arranged in fragmented habitat patches, the dimensions and separation of which vary considerably depending on the observational resolution. Reef associated species are hence fragmented into spatially discrete, seemingly isolated populations at any given spatial scale (m to 1000s km). The majority of coral reef fishes, like many other reef dwelling organisms, have a bipartite life history; while most species are relatively sedentary as adults, a pelagic larval phase theoretically offers the potential for extensive dispersal (Leis 2006). In most coastal marine species, larval dispersal is the primary demographic link between geographically separated populations. This process is known as population connectivity and more specifically refers to the exchange of individuals among local (sub)populations (Sale et al. 2005; Almany et al. 2009; Cowen & Sponaugle 2009). The extent to which coral reef fish populations are interconnected via larval dispersal has critical ramifications for metapopulation dynamics, resilience and the ability to recover from local disturbance (Kritzer & Sale 2004; Jones et al. 2007). Coral reef ecosystems worldwide are under increasing pressure from over-fishing, pollution, disease and climate change (Huges et al. 2003; Bellwood et al. 2004, Kennedy et al. 2013), putting corals and reef dwelling fishes at an elevated risk of extinction (Reynolds et al. 2005; Carpenter et al. 2008). Marine protected areas (MPAs) have been widely advocated as a means of strengthening the resilience of coastal habitats by acting as fisheries management and/or biodiversity conservation tools (Russ 2002; Jones et al. 2007; Botsford et al. 2009). The
effectiveness of MPAs as a buffer against anthropogenic impacts, however, depends fundamentally on the extent of population connectivity (Palumbi 2003). The degree to which reserve areas are demographically connected to each other and non-reserve areas is a key aspect in the design of MPA networks (i.e. size and spacing of individual reserves), yet efficient MPA network design is still hampered by a lack of understanding about larval dispersal (Sale et al. 2005; Almany et al. 2009; Botsford et al. 2009). So while measuring larval dispersal is regarded as one of the most challenging tasks in marine ecology, it is also a vital parameter for biodiversity conservation and fisheries management (Jones et al. 2005; Cowen et al. 2006; Hedgecock et al. 2007). While connectivity between populations may offer some degree of resilience to human impacts by the supply of recruits from un- and/or less impacted areas, the disruption of natural patterns of larval dispersal by over-exploitation, habitat loss and fragmentation may also be a central aspect in the potential collapse of coral reef fish populations in the near future (Hughes et al. 2005; Kritzer & Sale, 2006; Jones et al. 2009).

The last 15 years have witnessed a drastic increase in research on marine population connectivity, largely fueled by the need to incorporate information on larval dispersal into the design of MPA networks (Levin 2006; Jones et al. 2009; Hixon 2011). Despite this increasing interest and the resulting boost in knowledge, our understanding of larval dispersal in the marine environment is far from comprehensive (Thorrold et al. 2007; Jones et al. 2009). The following is a short and elementary review about some of the basic concepts of population connectivity in coral reef fishes, the approaches to study connectivity and the major recent advances in this field.
Concepts of connectivity:

Open or closed?

There has been a long-standing debate of whether metapopulation dynamics of coral reef fishes are mainly regulated by local self-recruitment or wide range dispersal and connectivity between subpopulations (e.g. Sale, 1991; Mora & Sale, 2002; Swearer et al. 2002; Warner & Cowen, 2002). Retention of larvae at the natal site would lead to “closed” populations that are mainly self-seeding, while “open” populations would export their larvae and receive their recruits mainly from external sources. Traditionally, coral reef fish larvae were believed to act like passive particles and that oceanic currents and long larval durations convey large-scale dispersal resulting in open populations (Caley et al. 1996; Roberts 1997), a paradigm supported by the demonstration of genetic panmixia over broad spatial scales (Shulman & Birmingham, 1995; Shulman 1998). Early models based on pelagic larval duration (PLD) coupled with prevailing current patterns suggested dispersal distances of 100s of km (Roberts 1997). Evidence emerging at the beginning of the new millennium, however, showed that coral reef fish larvae are far from being passive particles (Stobutzki & Bellwood 1994, 1997; Fisher & Bellwood 2002) and that larval behaviour (Tolimieri et al. 2000; Atema et al. 2002; Lecchini et al. 2005; Gerlach et al. 2007) and local- to mesoscale-current phenomena (Cowen et al. 2000; Paris & Cowen 2004) may result in high rates of local self-recruitment (Jones et al. 1999; Swearer et al. 1999; Warner & Cowen 2002). New field and laboratory techniques keep producing evidence for significant self-recruitment and dispersal in the range of meters to 10s of km (Thorrold et al. 2001; Jones et al. 2005; Cowen et al. 2006; Almany et al. 2007; Planes et al. 2009; Almany et al. 2013). Empirical estimates of self-
recruitment in coral reef fish typically range between 20-70%. It is now believed that most coral reef fish metapopulations exist in a continuum between open and closed states, which is regulated to varying levels by taxon and location-specific degrees of retention and connectivity (Jones et al. 2009). The potential capacity to replenish from either internal (retention) and/or external (immigration) sources would confer a higher level of resilience to individual populations than the existence as either completely closed or open populations (Almany et al. 2007). A thorough comprehension of connectivity and metapopulation dynamics, however, requires a deeper understanding of the dispersal process and insights into the “black box” of the larval stage (Kritzer & Sale, 2004; Sale et al. 2005; Jones et al. 2009). In Chapter 5 I present the first empirical evidence of a largely open reef fish population.

Evolutionary vs. demographic connectivity

By definition, population connectivity is the transfer of individuals among subpopulations, and hence entails both the replenishment of populations as well as the transfer of genes among them (Levin 2006; Cowen et al. 2007). It is thus useful to differentiate between two kinds of population connectivity and the terminology used throughout this thesis: evolutionary connectivity and demographic connectivity.

Evolutionary or genetic connectivity refers to the amount of intergenerational gene flow among populations determining the extent of genetic differentiation between them.

Demographic or ecological connectivity describes the exchange of individuals among populations that has a measurable influence on local demographics and population growth rates (Lowe & Allendorf 2010; Sale et al. 2010). The levels of exchange required
to impact populations demographically on ecological timescales is several orders of magnitude higher than those necessary to maintain genetic homogeneity between populations (one to two individuals per generation) over evolutionary timescales (Planes 2002; Cowen et al. 2007; Hedgecock et al. 2007). While evolutionary and demographic connectivity work on different temporal and spatial scales, an understanding of both processes is required for a full grasp on the patterns of connectivity in any given species (Lowe & Allendorf, 2010). In this thesis I combined approaches to quantify both types of connectivity (Chapter 4: evolutionary, Chapter 5: demographic).

Settlement versus reproductive connectivity

For dispersal to become influential on evolutionary and/or demographic scales, it is vital for settling larvae to actually survive to reproductive stage. If immigrants die before reproducing, dispersal is irrelevant for the long-term genetic composition of the population and its growth rates. It is hence also important to differentiate between settlement connectivity and reproductive connectivity (Hedgecock et al. 2007; Pineda et al. 2007; Lowe & Allendorf 2010). The majority of studies on reef fish connectivity in the past have concentrated on the former type, disregarding the high rates of post-settlement mortality that may significantly influence the dispersal outcome on a demographic scale. Recruits may vary in their post-settlement fitness depending on their source, genetic make-up and/or experiences during the larval stage. The influence of these factors on recruit survival is critical for management as it determines “realized” (reproductive) connectivity (Hamilton et al. 2008; Steneck et al. 2009). In Chapter 6 I
thoroughly discuss and conceptualize some of the prevalent theories concerning the factors affecting realized connectivity.

*The biophysical nature of dispersal*

The dispersal of minute propagules in a three-dimensional, highly stochastic environment is inherently a coupled biophysical process (Cowen *et al.* 2007; Werner *et al.* 2007). The interaction between local- and regional-scale oceanographic features with ontogenetically changing behavioural traits in developing larvae is a highly complex process that until now is not well understood (Cowen & Sponaugle, 2009).

While local current patterns may work to disperse or retain larvae within their natal environment, behavioural traits may work antagonistically or synergistically with these processes to result in varying spatial scales of dispersal from meters to 100s of km (Sponaugle *et al.* 2002; Cowen 2006). Physical forcing varies greatly between the open ocean and the near-shore environment. While information on oceanographic deep-water processes is relatively comprehensive, a lack of knowledge about the physical mechanisms in shallow-water environments is hampering more detailed modeling approaches of larval transport (Pineda et al 2007; Werner *et al.* 2007; Cowen & Sponaugle, 2009). Current flow in coastal regions exhibits considerable small-scale spatial and temporal complexity owing to frictional forces over complex topography, turbulence, tidal forcing, wind, water stratification etc. (Gawarkiewicz *et al.* 2007). There is a strong bias in physical connectivity studies towards explanations based on meso-scale processes that can be measured remotely (e.g. satellite oceanography), while ignoring
complex but potentially vital small-scale mechanisms (Pineda et al. 2007). There is a crucial need to assess the near-shore processes that influence larval transport combined with rigorous field testing of model predictions (Pineda et al. 2007; Werner et al. 2007). In Chapter 5 I compare empirical estimates of larval dispersal in a coastal reef system with the predictions of a biophysical dispersal model.

As discussed earlier, physical processes alone do not determine connectivity in coral reef fishes, but internal and external biological attributes during the pelagic larval stage have also been shown to play important roles in the dispersal process (Leis 2002; Armsworth & Roughgarden 2005; Cowen & Sponaugle 2009). Species-specific attributes like pelagic larval duration (PLD), ontogeny, vertical and horizontal swimming behaviours, as well as breeding modes and locations may all act to influence larval trajectories (Leis 2002; Cowen et al. 2000; Cowen 2006; Pineda et al. 2007). While spawning mode (broadcast or brooding) and PLD seem to have comparatively little influence on rates of connectivity (Almany et al. 2007; Jones et al. 2009; Weersing & Toonen 2009), oriented swimming behaviour in response to settlement cues and/or vertically structured currents has been shown to significantly influence modeled dispersal patterns (Kingsford et al. 2002; Leis 2002; Gerlach et al. 2007; Leis 2007).

This complex interplay between biological and physical factors results in highly variable dispersal distances and trajectories within and between reef fish species probably to a large degree dictated by the availability and spacing of suitable settlement habitat (Jones et al. 2009). Overall, despite increasing interest and growing understanding about the field of population connectivity, large gaps of knowledge remain and dispersal
patterns are still a key uncertainty in coral reef management and MPA design (Botsford et al. 2009).

**Approaches to study connectivity:**

The minute size and spatial spread of pelagic larvae in concert with the extremely high fecundity of most marine species and substantial mortality rates at early life stages make direct tracking of reef fish larvae essentially impossible (Thorrold et al. 2002; Levin 2006). Moreover, the wide range of larval behaviours in combination with spatially variable physical mechanisms in the sea makes dispersal an extremely complex process to measure (Pineda et al. 2007). Nevertheless, a number approaches exist to study connectivity on different temporal and spatial scales. Coral reef fish studies in particular have been at the vanguard of the development of novel research approaches and the refinement of existing techniques during the last 15 years (Jones et al. 2009; Sale et al. 2010). Approaches can generally be divided into direct and indirect methods.

*Direct approaches* include the analysis of island endemism (Robertson 2001) and *in situ* observations of larval behaviour and trajectories of individual propagules (reviewed in Leis 2002, 2007). A more recent approach is the artificial tagging of larvae prior to hatching and recapture after settlement (reviewed in Thorrold et al. 2002, 2007; MacMahon et al. 2013). This method involves the marking of otoliths (ear bones) in larvae by the immersion of eggs in tetracycline solution (Jones et al. 1999; Jones et al. 2005), or via maternal transmission of enriched stable isotopes injected into mature females prior to egg production (Thorrold et al. 2006; Almany et al. 2007). Tagging is
labor intensive and requires the marking of significant proportions of the total larval population to overcome the extensive mortality rates and diffusion of larvae in the pelagic, yet can yield direct empirical evidence of the scale of dispersal (Thorrold et al. 2007; Cowen & Sponaugle 2009).

Indirect approaches include phylogeography (reviewed in Rocha et al. 2007) and the analysis of the elemental chemistry of otoliths (geochemical signatures) that can be used to discriminate between sources of larvae (Swearer et al. 1999; Thorrold et al. 2001, 2007; Hamilton et al. 2008; Ruttenberg et al. 2008). The latter approach eliminates the problem of low recapture rates because every single larva is naturally marked. However, significant uncertainties about the identification of elemental tags can arise in the absence of clear long-term geochemical signatures of local water bodies, entailing the need to sample many potential source locations (Thorrold et al. 2007; Cowen & Sponaugle 2009).

The most widely used method to make inferences about larval dispersal and connectivity in coral reef fishes (and other marine organisms) is the application of genetic markers (reviewed in Planes 2002; Hedgecock et al. 2007; Hellberg 2007). Genetic tools in turn divide into indirect and direct methods: population genetics and parentage and assignment tests respectively.

*Population genetics*

The use of population genetic structure to infer rates of population connectivity has increased rapidly over recent decades owing to the proliferation of available markers and
the improvement of laboratory protocols for the analysis of genetic data (Hellberg 2007; Cowen & Sponaugle 2009). Classic population genetic analysis seeks to measure the amount of gene flow between subpopulations. If populations are demographically isolated from each other, restricted gene flow eventually leads to genotypic differentiation between these populations due to natural selection and random genetic drift. On the other hand, if populations are connected by regular gene flow, they will remain genetically homogeneous. These genetic differences are measured as differences in allele-frequencies at specified loci (the location of a gene on the chromosome). The most widely used measure of population differentiation/gene flow is $F_{ST}$ (Wright 1943), the standardized variance in allele frequency, which represents the proportion of total genetic variation partitioned among subpopulations (Hellberg 2007; Lowe & Allendorf 2010). Under the assumption of genetic equilibrium (Slatkin 1993), $F_{ST}$ can be used to indirectly calculate the number of individuals exchanged between populations per generation. The underlying Island Model (Wright 1943) is based on many unrealistic assumptions however, and $F_{ST}$ needs to be interpreted with caution, especially in high gene flow scenarios (Waples 1998; Lowe & Allendorf 2010).

Theoretically, even small amounts of gene flow are sufficient to maintain genetic homogeneity between separated populations (Planes 2002). Ten migrants per generation are enough to forestall all but minor allele-frequency changes, and even only one successful migrant per generation can prevent the accumulation of any large genetic differences (Slatkin 1987). Nevertheless, the literature includes a surprisingly large number of studies that have found high genetic differentiation despite high dispersal potential (e.g. Lacson 1994; Lacson & Clark 1995; Planes 1993; Planes et al. 1998;
Shulman & Bermingham 1995; Planes & Fauvelot 2002; Taylor & Hellberg 2003, 2005; Gerlach et al. 2007). In the absence of obvious physical barriers, this fine-scale genetic differentiation is mostly invoked to be a result of larval retention by physical factors and/or natal homing (Taylor & Hellberg 2003, 2005; Gerlach et al. 2007). Some studies found both low and high differentiation in different species at similar spatial scales (e.g. Doherty et al. 1995; Planes et al. 1998), adding to the notion that PLD is a good predictor for dispersal distance (Bohonak 1999; Siegel et al. 2003). Overall, it seems however that spatial differentiation is more important for genetic differentiation than might be expected from the length of larval duration alone (Bay et al. 2006; Planes 2002; Weersing & Toonen 2009).

In general, $F_{ST}$ in marine fishes is comparatively low, indicating commonly high gene flow among segregated populations (Ward et al. 1994; Waples 1998). The majority of studies on coral reef fish population structure accordingly have found low genetic differentiation, and in most cases have inferred high dispersal and panmixia among subpopulations (e.g. Shulman & Bermingham 1995; Froukh & Kochzius 2007; Horne et al. 2008; Hepburn et al. 2009; Purcell et al. 2009; Jones et al. 2010; Salas et al. 2010; Visram et al. 2010). Low values of genetic differentiation, however, do not necessarily imply high rates of connectivity because genetic equilibrium cannot be assumed for every population (Palumbi 2003; Hedgecock et al. 2007). Additionally, the actual genetic signal to sampling noise ratio may become exceedingly low in high gene flow data, making standard estimates of migration increasingly inaccurate (Waples 1998). Fortunately, new approaches based on maximum likelihood and/or Bayesian inference are now available that bypass many of the assumptions inherent in population genetic models.
Another tool to help address the problem of high gene flow is the isolation by distance (IBD) model (Wright, 1978) that allows for the careful measurement of the buildup of genetic differences as a function of distance (Waples 1998; Palumbi et al. 2003). More specifically, IBD describes the increase in genetic differentiation at neutral loci with increasing geographic distance (Slatkin, 1993). An increasing number of studies now apply IBD models to their population genetic data (e.g. Froukh & Kochzius, 2007; Hepburn et al. 2009; Jones et al. 2010; Villegas-Sanchez et al. 2010; Visram et al. 2010), as this model seems to be more suitable for most marine populations than the Island Model and to some extent mitigates the problem of sampling noise in high gene flow systems (Waples 1998; Palumbi 2003). IBD represents a balance between dispersal and genetic drift as high differentiation between spatially distant populations might either result from low migration or low effective densities (or a combination thereof). Using estimates of local effective densities, IBD can hence be used to infer dispersal distances (Puebla et al. 2009; Pinsky et al. 2010).

In the end, indirect population genetic approaches like the mentioned stepping-stone model have proven crucial to assess the evolutionary consequences of dispersal, however, they are unable to differentiate between historical and contemporary gene flow (Hedgecock et al. 2007; Lowe & Allendorf, 2010; Sale et al. 2010). Patterns of genetic connectivity should hence not be misinterpreted to make unsubstantiated inferences about demographic connectivity at ecological timescales.
Parentage analysis

Recent advantages in molecular tools and statistical analysis have led to the development of genetic methods to produce direct estimates of connectivity in marine systems (Manel et al. 2003). These approaches center on the assignment of individuals to populations of origin (assignment tests) or to their parents (parentage analysis) (Manel et al. 2005). In situations of high gene flow, parentage analysis seems to produce more accurate estimates of connectivity than assignment tests (at least at small spatial scales) (Saenz-Agudelo et al. 2009) and might hence be the more suitable method for most marine species. Parentage analysis assigns individuals to the most likely parents from a pool of potential parents most often using maximum-likelihood analysis (Jones & Ardren 2003). While recently developed methods based on Bayesian approaches have overcome many of the intrinsic assumptions (Manel et al. 2005; Christie et al. 2010a, b), the accuracy of parent allocation declines drastically with incompletely sampled adult populations; i.e., parentage analysis works best if all potential parents can be sampled (Hedgecock et al. 2007; Jones et al. 2009). Conceptually this approach is similar to other direct approaches like artificial tagging (Hedgecock et al. 2007). Jones et al. (2005) combined parentage analysis and tetracycline tagging to directly estimated self-recruitment in an anemonefish and found remarkably similar results with both methods (c. 30% natal homing), essentially providing a validation of parentage as a reliable technique.

Parentage studies have so far provided great insight into the dispersal process of coral reef fishes, proving that substantial amounts of recruitment in small MPAs may be spawned locally, but also that larvae may travel distance in excess of 30 km between MPAs (e.g. Planes et al. 2009; Almany et al. 2013).
Overall, indirect and direct genetic approaches operate on opposite ends of the temporal spectrum. While indirect methods estimate connectivity over evolutionary timescales, direct techniques measure dispersal and retention over a single or a few generations (Hedgecock et al. 2007). Integrating both approaches holds promise to be a more suitable strategy to elucidate the dispersal kernel than any one technique alone (Hedgecock 2010). In this thesis I combined population genetic approaches on large, Red Sea wide, spatial scales (Chapter 4) with parentage analysis in a more localized reef system (Chapter 5).

Coupled biophysical models

In addition to determining the shape of genetic population structure, it is important to verify the processes that actually lead to the observed patterns (Kool et al. 2010). As mentioned above, patterns of larval dispersal are created by a complex interplay between physical (current flow) and biological (e.g. larval behaviour, mortality) mechanisms. Field studies are always snapshots in space and time and will never be able to cover the entire range of spatial and temporal variability involved in larval transport. Coupled biophysical models have hence become a critical tool to capture and quantify the processes that drive marine population connectivity (Cowen 2006; Werner et al. 2007; Paris et al. 2013).

Physical models of ocean circulation are becoming increasingly sophisticated and are available at different spatial resolutions and levels of complexity (Cowen et al. 2006; Werner et al. 2007; Paris et al. 2013). While early models of dispersal were relying solely
on physical processes, it is now apparent that biological parameters are of equal importance to model larval trajectories (Leis 2007). Spatially explicit individual-based models (IBM) nested in Lagrangian stochastic schemes have now become a key tool to investigate larval fish movements in complex coral reef environments (e.g. Cowen et al. 2006; Paris et al. 2007; Paris et al. 2013). These models allow for active particle tracking simulations informed by physical (e.g. currents, diffusion) as well as biological (e.g. vertical migration, mortality) parameters. Several studies have attempted to couple dynamic physical oceanographic data with the biological parameters involved in larval dispersal, varying in their model complexity and the covered spatial scale (e.g. Cowen et al. 2000, 2003, 2006; James et al. 2002; Paris et al. 2005, 2007; Treml et al. 2008; Foster et al. 2012). Collectively these studies demonstrate the ability of hydrodynamic models to predict larval dispersal trajectories and patterns of connectivity between subpopulations at various scales. In Chapter 5 we implement the first biophysical model of larval dispersal in the Red Sea.

Despite their high level of sophistication, modern connectivity models are still hampered by imprecise parameterizations of input variables (Paris et al. 2007; Werner et al. 2007; Simons et al. 2013). Crucial questions remain about the influence of larval behaviour in the dispersal process. How does behaviour weigh against physical forcing? How does active larval movement (vertical and horizontal) affect passive transport by currents and tides? While larval behaviour is clearly significant in coral reef fish dispersal, its relative importance remains largely unknown and difficult to measure (Sale 2004; Leis 2007; Treml et al. 2008).
**Anemonefish as model organisms:**

There are 29 known species of anemone- or clownfish (Amphiprioninae), all but one belonging to the genus *Amphiprion*. Clownfish usually form small groups of one breeding pair together with one to five non-breeders. Once the dominant female dies, the largest male changes sex to become the female, a system known as protandrous hermaphroditism (Fricke & Fricke, 1977; Buston 2003). Anemonefish live in facultative mutual symbiosis with sea anemones, a microhabitat they rarely stray farther than a few meters away from during their entire adult life. Spawning usually occurs around the full moon when around 500-800 eggs are deposited and fertilized on a flat surface below the anemone (Maroz & Fishelson 1997). Larvae hatch after about one week and PLD lasts around 8-12 days. Several species of anemonefish have been shown to be capable swimmers during the pelagic stage (Fisher et al. 2000, 2005) and late stage individuals may utilize chemical cues emitted by their host anemones to locate settlement habitats (Elliott et al. 1995; Munday et al. 2009). Anemonefish may hence exhibit significant behavioural control over their dispersal trajectories, at least at later stages of development.

Genetic research on *Amphiprion clarkii* revealed high gene flow over distances of up to 1000 km on the Great Barrier Reef (Doherty et al. 1995), indicating extremely high dispersal potential for this genus. Recent studies using parentage analysis (Jones et al. 2005; Planes et al. 2009; Saenz-Agudelo et al. 2011; Berumen et al. 2012) and artificial tagging (Almany et al. 2007), on the other hand, have produced high estimates of self-recruitment (20-60%) in *A. polymnus* and *A. percula*. The genus hence seems to exhibit variable patterns of connectivity characterized by a broad dispersal kernel. To assess the...
full range of dispersal patterns in an *Amphiprion* species it is thus vital to apply and combine methods covering multiple spatial and evolutionary scales.

In the Red Sea there is only one species of clownfish, the two-banded anemonefish, *Amphiprion bicinctus* (Rüppel, 1830) (Fig. 1). *A. bicinctus* are common inhabitants of shallow coastal reef areas and live in symbiosis with the anemone species, *Heteractis magnifica*, *Entacmea quadricolor*, *Stichodactyla daddoni* and to a lesser extent with *Heteractis crispa* and *Heteractis aurora* in depths between 0.5 and 40m.

![Adult, male individual of the two-band anemonefish, *Amphiprion bicinctus*](image)

**Figure 1:** Adult, male individual of the two-band anemonefish, *Amphiprion bicinctus*

Overall, anemonefish are ideal model organisms to study connectivity. Their strong habitat association greatly facilitates research in that (1) it eliminates the influence of spawning migrations, i.e. spawning and potential settlement locations are known in advance, (2) the same individuals can be re-sampled consecutively, (3) for parentage analysis, typical densities make it possible to sample the majority of potential parents at a given location (see Jones *et al.* 2005; Planes *et al.* 2009).
Connectivity research in the Red Sea:

Compared to other parts of the world, research on population connectivity in the Red Sea has been scarce (Berumen et al. 2013). The majority of studies in the Red Sea have been conducted in the Gulf of Aqaba (mostly Israel and Jordan) and coral reef research in any other region, especially along the Saudi Arabian coastline has so far been extremely limited.

In his dissertation, Froukh (2007) analyzed the population genetic structure of three coral reef fishes, the endemic fourline wrasse, *Labricus quadrilineatus*, the blue-green chromis, *Chromis viridis*, and the lyretail anthias, *Pseudanthias squamipinnis*, along the Saudi Arabian coastline up to the Gulf of Aqaba (Jordan). He found low genetic differentiation for all three species throughout the Red Sea, yet discovered significant variation between the northern and central/southern populations of *L. quadrilineatus* and between the Gulf of Aqaba and the Red Sea proper in *P. squamipinnis*. Migration analysis revealed that general larval movement patterns may differ between the study species (northward for *C. viridis* and *L. quadrilineatus*, southward for *P. squamipinnis*). These results suggest that different species, although sharing the same habitat and similar life styles, may vary substantially in their larval dispersal patterns. Although historical patterns of gene flow are undoubtedly important (Planes 2006; Froukh 2007), a lack of information on oceanographic processes in the Red Sea has so far been hampering the interpretation of such population genetic data.

Further population genetic studies have shown panmixia in the lionfish *Pterois miles* throughout the northern Red Sea and the Gulf of Aqaba (Kochzius & Blohm 2005) and in
several Siganidae species between the northern Red Sea and the Mediterranean (via Lessepsian migration) (Bonhomme et al. 2003; Hassan & Bonhomme 2005; Azzurro et al. 2006). More recently, Ben-Tzvi et al. (2008) attempted to use biogeochemical otolith signatures to determine dispersal routes in C. viridis within the northern Gulf of Aqaba. Their results were unclear, yet imply heterogenic larval supply from external sources (most likely the Saudi Arabian and Egyptian coast).

Clearly, patterns of connectivity between reef fish populations in the Red Sea are far from being resolved and further information, especially concerning small-scale processes, is needed to make informed management decisions. Red Sea coral ecosystems are being increasingly put under stress from anthropogenic impacts and networks of MPAs along the Saudi Arabian coastline have been proposed to mitigate some of these pressures (PERSGA, 2010). MPAs should be designed upon a sound understanding of population connectivity (Botsford et al. 2009), clearly demanding for more research in this area within the Red Sea basin. Here I combined several approaches (i.e. population genetics, parentage analysis, biophysical models) at different spatial scales to study the patterns of larval dispersal in a coral reef fish, A. bicinctus, in the Red Sea.
References:


CHAPTER 2: REVIEW

On the mechanisms of marine larval connectivity

Introduction:

Research relating to larval dispersal, retention and connectivity on coral reefs has proliferated drastically over the past two decades (Jones et al. 2009; Hixon 2011) and numerous research groups worldwide are presently trying to understand and forecast the fate of pelagic larvae of different taxa using various technical approaches (reviewed in Leis et al. 2011a). This remarkable increase in connectivity research is not necessarily due to a recent major shift in research priorities, but more likely to a growing recognition of the importance of understanding connectivity and the ability of many types of research to inform basic questions related to this topic.

Recent advances in molecular biology, otolith microchemistry and the use of artificial larval markers combined with novel statistical methods have provided the necessary tools for empirically testing larval dispersal distances, population connectivity, and export effects of marine reserves (e.g. Thorrold et al. 2007; Planes et al. 2009; Leis et al. 2011a; Saenz-Agudelo et al. 2011; Harrison et al. 2012). At the same time, coupled biophysical dispersal models are rapidly becoming more sophisticated, integrating biological parameters and oceanographic patterns to feed active particle tracking routines (e.g. Paris et al. 2007; Siegel et al. 2008). In theory, predictions of these models can now be validated against such empirical estimates of dispersal (see Foster et al. 2012, Sponaugle et al. 2012), a process we will certainly see applied frequently in the near future.
However, while we can track larvae from point A to point B (at very much restricted spatial and temporal scales), we still lack insight into the black box of the pelagic phase. In other words, while we may be able to better measure and predict individual dispersal pathways, we are far from reliably simulating larval transport (*sensu* Pineda *et al.* 2007). The factors that drive larvae from reef A to reef B (or back to reef A), and the degree to which larval behaviour can offset or augment physical forcing are still poorly understood.

Marine larval dispersal is an exceedingly complex topic involving countless interacting factors at any given scale. We are looking at an amalgamation of chemical, physical and biological processes driving minute yet actively behaving and developing larvae in a vast and highly stochastic environment. The sheer plethora of involved elements makes it seem utopic to ever fathom the process of larval connectivity in its entirety; nor does it seem likely that we will ever be able to reliably predict individual trajectories for all locations at all times. Yet, as methodologies advance and research efforts increase, the bigger picture may eventually evolve. In this chapter I describe and discuss emerging patterns of underlying mechanisms and processes that appear to influence connectivity on different scales and point out gaps in information that might be filled by related research efforts.

Factors that influence dispersal, connectivity and the replenishment of populations are varied and they act in different ways and on different scales to influence these processes. The factors that influence larval dispersal and connectivity act over temporal scales ranging from seconds (e.g. behavioural decisions), over days (e.g. vertical migration, larval ontogeny), to months and years (e.g. seasonality, ENSO). The factors determining connectivity also vary over a range of spatial scales from mm (e.g. micro-habitat
preferences) to 100s km (e.g. long-shore transport). Moreover, dispersal kernels (the probability density function of dispersal distances) are probably highly site- and/or species-specific, driven by a vast number of interrelated biotic and abiotic, direct and indirect variables acting across the aforementioned spatial and temporal scales. In addition, virtually all of the involved factors shaping connectivity exhibit some sort of temporal variability and the patterns of one season might change to the next (Hogan et al. 2012; Berumen et al. 2012).

Exactly how all these different factors interact and balance against each other with regard to their impact on larval transport is as yet effectively unknown (Pineda et al. 2009). However, there is a great need to be able to generalize about the scale and magnitude of connectivity. A number of variables have been identified as potential drivers of dispersal and some have been employed as proxies for dispersal or connectivity in lieu of direct estimates (e.g. PLD, see below). In the simplest of worlds, a handful of basic characteristics (e.g., vertical migration of larvae, geography of the reef matrix, reef-scale hydrodynamics) could emerge as consistent predictors of dispersal and/or connectivity (e.g. see Pinsky et al. 2012). To date though, such general predictors have not been identified across taxa and interactions among factors seem to be too complex to simplify the underlying patterns. Theoretically, however, isolating key variables or combinations of variables would allow for generalizations for a given species/location if these variables could be estimated reliably and incorporated dependably into dispersal models.

In this context, numerous studies may contribute valuable information to the bigger picture of connectivity, even if their focus is not on connectivity directly. The following
points are meant to serve as examples of how a broad range of existing and future research could be aligned to help address the gaps in our current understanding of the drivers of dispersal. Clearly no single factor here is adequate to accurately predict patterns of dispersal or connectivity, but the integration of these data will greatly strengthen model predictions and may profoundly amplify the impact of science on management decisions.

**Species-level traits**

*Pelagic larval duration*

Dispersal among reefs is mediated primarily by the dispersal of pelagic larvae via the interaction between behavioural traits and ocean currents (Sale 1980; Kritzer & Sale 2004). Thus we can reasonably presume that the duration of the larval period will have a significant effect on the dispersal capabilities of a given species (Lester *et al.* 2007). It has been shown that the pelagic larval duration (PLD) is related to genetic estimates of historical dispersal ($F_{ST}$) among populations in a broad range of marine taxa including algae, invertebrates and fishes (Bradbury *et al.* 2008) and among species of reef fishes specifically (Doherty *et al.* 1995). However, this does not seem to be a general rule across all taxa and evidence is mounting that PLD is neither a good predictor of evolutionary connectivity (Weersing & Toonen 2009) nor range size (Mora *et al.* 2012). Bowen *et al.* (2006) for instance showed that two related holocentrid species with quite different PLDs showed contradictory levels of historical dispersal in the Atlantic, indicating the ineffectiveness of PLD as a predictor for the levels of gene flow on larger scales. While
the debate about PLD as a reflection of dispersal scales on a gene flow level continues (e.g. Riginos et al. 2011; Selkoe & Toonen, 2011; Faurby & Barber, 2012), it more likely sets an upper threshold for the dispersal distances achievable, but exerts less influence on local- or meso-scale connectivity. Evidence for this idea comes from remarkably similar estimates of self-recruitment for species with quite different larval durations (Almany et al. 2007; Berumen et al. 2012). Similarly, despite larval durations of 6-8 weeks, a gastropod species rarely exceeded dispersal distances of 5 km in a set-up of artificial settlement stations along the coast of northern New Zealand, indicating that, while long distance dispersal is possible, even species with long PLDs show a dispersal kernel skewed towards the natal environment (Salinas-de-Léon et al. 2012).

Larval behaviours

Early models of coral reef fish population dynamics assumed larval dispersal to be a passive process, the rate of which was a product of current transport and larval duration (e.g. Roberts 1997). Over the past two decades, however, numerous studies have demonstrated that larvae may indirectly and/or directly influence their own dispersal through various mechanical and sensory mechanisms (reviewed in Montgomery et al. 2001; Leis 2006; Leis et al. 2011b). Coral reef fish larvae in particular were shown to be capable swimmers (e.g. Stobutzki & Bellwood 1997; Fisher et al. 2005; Fisher & Leis 2009) with sensory abilities that enable active orientation in the water column (e.g. Leis et al. 1996; Lecchini et al. 2005; Simpson et al. 2005; Montgomery et al. 2006; Dixon et al. 2008; Irisson et al. 2009). Both larvae of reef fish and corals were shown to respond
positively to the chemical (Dixon et al. 2008; Gleason et al. 2009) and auditory (Simpson et al. 2005; Vermeij et al. 2010) cues produced by potential settlement reefs. Rising levels of ocean acidification on the other hand were demonstrated to potentially offset olfactory (Munday et al. 2009) and auditory (Simpson et al. 2011) orientation in reef fish, as well as chemical settlement responses in corals (Doropoulos et al. 2012), possibly resulting in significantly altered dispersal and survival patterns for impacted species.

Active behaviours like diel or ontogenetic vertical migrations in reef fish larvae were shown to notably affect connectivity, causing retention in onshore moving currents, thus restricting dispersal (Taylor & Hellberg 2003; Paris & Cowen 2004; Huebert et al. 2011). Horizontally stratified currents for instance may allow larvae to vertically probe the water column and then swim against the current carrying a favourable olfactory cue (Myrberg & Fuiman 2002). Alternatively, onshore currents may actively be chosen to “hitch a ride” back to the reef, using sound cues as orientation aid. Recent research indicates that vertical migration patterns might be strongly taxon and/or species specific (Huebert et al. 2011) and there is a critical gap in data integrating vertical larval movements with stratified current patterns.

Yet, even provided with better empirical data, we would still lack an understanding of the causation behind observed patterns measures using empirical approaches. Which factors for instance regulate individual decisions about vertical migrations, foraging, and predator avoidance considerations, or the chance to navigate using stratified currents?

While many modern dispersal models incorporate dynamic behavioural modules, parameters are typically standardized throughout the cohort, disregarding individual, rule-based decisions (sensu Fiksen et al. 2007). Whether behavioural decisions are
deliberate or genetically inherited, they are triggered by internal (physiological) and external (stimulus-driven) conditions encountered during dispersal. We still know very little about the “reasoning” behind larval choices, impeding the development of rule-based modules for individual-based models.

While most behavioural work has so far focussed on coral reef fish larvae, substantial literature is also available on oriented swimming in crustaceans (e.g. Jeffs et al. 2003, 2005; Montgomery et al. 2006; Radford et al. 2007; Stanley et al. 2010) and sponges (e.g. Maldonado 2006; Mariani et al. 2006; Abdul Wahab et al. 2011). Yet comparatively little information exists for other invertebrate taxa. Most work on corals has so far concentrated on micro-scale settlement preferences, specifically determining the role of crustose coralline algae (e.g. Harrington et al. 2004; Ritson-Williams et al. 2010; but see Raimondi & Morse 2000; Stake & Sammarco 2003). There remains a definitive need for more research to address behavioural patterns and orientation abilities in a broader range of taxa.

Importantly also, most existing studies on orientation behaviour were implemented in artificial settings such as flumes or aquarium tanks that do not accurately reflect the plethora of sensory inputs experienced by larvae in the field (Pineda et al. 2009). It is now clear that olfaction (Atema et al. 2002; Gerlach et al. 2007, Dixson et al. 2011), sound (Tolimieri et al. 2000; Montgomery et al. 2006) and/or vision (Lecchini et al. 2005) have the theoretical potential to serve as navigation aids, and biophysical models predict increased settlement success with larger sensory detection distances (Staaterman et al. 2012). The scales at which these stimuli may interact in situ, however, remain virtually unknown.
Larval ontogeny

Perhaps as important as larval behaviour itself are the ontogenetic onset and the development of such behaviours during the pelagic phase (Leis & McCormick 2002; Clark et al. 2005; Miller & Kendall Jr 2009). The incorporation of active larval behaviour in dispersal simulations resulted in significantly different model predictions as compared to passive drift (Paris et al. 2007; Staaterman et al. 2012). It should hence be critical at which point larvae begin to exhibit such behaviours and how they develop during ontogeny. While most work on the morphology of sensory organs (e.g., Lara 2008; Wright et al. 2010) and navigation towards sensory cues (e.g. Tolimieri et al. 2000; Atema et al. 2002; Lecchini et al. 2005; Gerlach et al. 2007) so far has been done on settlement stage larvae only, a growing recognition of the importance of ontogenetic development has recently led to an increasing number of publications focusing on different features of this aspect. The emerging picture is one of significant changes in orientation abilities during larval development. Moreover, this development shows strong inter- and even intra-specific variations (reviewed in Leis 2010).

The development of swimming capabilities in reef fish larvae for instance was shown to exhibit marked species-specific patterns. While critical swimming speeds increase during development (related to size), the rate at which this happens appears to differ between families and even species within families (Fisher et al. 2000; Clark et al. 2005; Leis et al. 2007). Especially in species with better swimming abilities, transport trajectories will differ substantially depending on the amount of time elapsing before larvae begin actively orientating in the water column.
Moreover, there might be ontogenetic changes in sensory responses over time, influencing swimming trajectories. Dixon et al. (2011) found that larvae of two anemonefish species were able to utilize olfactory cues early in their development, but that preferences changed during the 11-day pelagic stage. The authors concluded that while at earlier stages larvae may be inclined to disperse away from the reef (towards pelagic waters), they might become attracted and navigate towards settlement cues of appropriate habitats at later stages.

Hearing abilities in reef fish larvae were also shown to be present early in development, improving with growth (Arvedlund & Kavanagh, 2009). While auditory development seems to differ greatly among species (Wright et al. 2011), ontogenetic shifts in response to stimuli as observed by Dixon and co-workers (2011) may also occur, guiding larvae in different directions depending on their developmental stage. Behavioural shifts like these may profoundly affect model predictions and should carefully be incorporated into the parameter set.

Even such seemingly straightforward parameters as settlement competency may exhibit subtle intricacies. Connolly & Baird (2010) parameterized a simple dispersal model with experimentally derived estimates of settlement competencies in five coral species. They found that within-cohort variations in competency dynamics had profound effects on their model estimates of self-recruitment and long-distance dispersal. The morphological development of sensory abilities, coupled with the ontogenetic onset of active behavior thus seems to be critical in determining dispersal trajectories and any effective biophysical model will need to be parameterized accordingly. With regard to recent findings, research on a wide suite of taxa and species will be necessary to account
for inter-specific differences in sensory and physical developments. There is a pressing need for innovative approaches to test these behaviors in a more natural setting as responses in experimental aquaria may not reflect natural behaviors in the wild (e.g. Irrison et al. 2009).

**Larval Mortality**

Larval survival is a highly important component of connectivity, yet patterns of mortality during the pelagic phase are still poorly understood. The spatial scale of connectivity will be, in part influenced by the likelihood of larvae surviving transport and settling upon a reef. Conceptually, the odds of a larva making it to another reef should decrease as a function of distance between reefs. The shape of the mortality function (or mortality schedule) is hence expected to influence the scale of demographic connectivity (Graham et al. 2008; Steneck et al. 2009) and as such represents a vital parameter for biophysical connectivity models (Werner et al. 2007). Yet, estimates of larval mortality are rare and mostly inferred from laboratory data (Pineda et al. 2009). Reliable data on ontogenetic changes in mortality are virtually absent and most dispersal models use the idealized and unrealistic assumption of constant mortality during the larval phase (but see Connolly & Baird 2010). More work on survival probabilities over a broad range of taxa is critically needed to enhance parameter inputs for realistic dispersal models.

Additionally, mortality is clearly not only stochastic within, but also between cohorts and survival probabilities will vary from one recruitment pulse to the next according to varying intrinsic and external biotic and abiotic factors (Sale 2004; Siegel et al. 2008).
Recent work indicates that larval mortality may also be impacted by climate change. While ocean acidification did not have a direct effect on the survival of larvae of the spiny damselfish *Acanthochromis polyacanthus* (Munday *et al.* 2011), it was demonstrated that clownfish larvae lose their ability to use olfactory cues for the location of suitable settlement sites (Munday *et al.* 2009) as well as for predator avoidance (Dixon *et al.* 2010). The loss of predator recognition was shown to result in increased levels of mortality (up to 9-fold) during settlement (Munday *et al.* 2010), indicating that such behavioural alterations may have far-reaching implications for connectivity patterns and population replenishment. Tolerances to acidification, however, may vary significantly between even closely related species (Ferrari *et al.* 2011) and more research over a broad range of taxa will be pivotal to address individual species responses to the combined effect of global warming and ocean acidification on larval survival.

*Larval quality*

Larvae of reef organisms often exhibit intraspecific variations in phenotypes (e.g. size, growth rates) that are believed to affect subsequent juvenile performance and potentially carry over to later life stages (Shima & Findley 2002; McCormick & Hoey 2004; Shima & Swearer 2010; Smith & Shima 2011). While post-settlement processes are not the focus in this chapter, larval quality may also impact on survival rates, recruitment success and/or dispersal patterns (Wilson & Meekan 2002; Sponaugle *et al.* 2006, see Chapter 6). Faster growth rates in developing larvae for instance were shown to translate into higher survival rates in the plankton (Wilson & Meekan 2002). Intraspecific variations in larval
phenotype may hence profoundly influence settlement success and mediate patterns of
connectivity in coral reef metapopulations (Shima & Swearer 2009).

Variance in phenotype or larval quality has been mostly attributed either to conditions
encountered in the pelagic environment (see below) or to environmentally induced
parental effects. Resource limitations in spawning mothers for instance might induce
phenotypic variations in larval conditions, potentially influencing survival and/or
recruitment success (McCormick 2006; Gagliano & McCormick, 2007). Larval quality
might hence to some extent be a function of the natal environment (Shima & Swearer,
2009). Where resources (e.g. food, shelter, habitat) are patchily distributed in a reef
matrix, parental effects on dispersing larvae may hence have far-reaching consequences
for recruitment dynamics and meta-population connectivity (see Chapter 6 for a more in-
dept discussion).

Genotype (individual differences)

Evidence is mounting from terrestrial systems that there may be genetic predispositions
for differentiated phenotypes related to dispersal (Fitzpatrick et al. 2005). Theoretical
models have shown that individuals at the tail of the dispersal kernel (i.e. long distance
dispersal) can differ in dispersal-associated genotypes compared to individuals closer to
the start of the curve (i.e. short distance dispersal) (e.g. Phillips et al. 2008). Haag et al.
(2005) managed to identify a specific allele (‘candidate gene’, sensu Fitzpatrick et al.
2005) in butterflies responsible for flight metabolic rate, the frequency of which was
highly correlated with dispersal rate. Similarly, Bitume et al. (2011) observed heritable
maternal effects on inter-generational dispersal behaviour in spider mites, where under low density conditions, offspring of mothers with long dispersal distances also traveled far.

If similar selective processes apply for dispersal patterns in reef organisms, we may observe genetic predispositions for self-recruitment and/or long distance distribution, which may result from genotypic and/or phenotypic conditions related to dispersal (e.g. metabolic rate, fin size). Additionally, heritable behavioural traits may have a strong influence on individual dispersal trajectories when discrete responses to external stimuli differ between genotypes. Even minor individual decisions concerning vertical positioning in the water column for instance may cause entirely distinct larval trajectories, entailing potentially large effects on growth, survival and eventually on the settlement location (Fiksen et al. 2007).

Adult behaviour

Larval dispersal is predominantly affected by processes that occur during the larval phase, however, the spawning behaviours of adults may also significantly affect dispersal potential. Owing to chaotic hydrodynamics around topographically complex reef environments, minute differences in the release point of a larva may have considerable effects on its trajectory (Gawarkievicz et al. 2007). Likewise, the timing of the release of eggs and larvae is thought to affect the dispersal of propagules from the reef. Lunar patterns of spawning and larval release for instance are thought to coincide with the tidal amplitude cycle, expectedly influencing the degree of early dispersal (Sponaugle et al.)
Species exhibiting spawning migrations and/or synchronously timed spawning may hence differ substantially in their dispersal patterns from species without such adult behaviours, even if larval behavioural (e.g. vertical migration) and life-history characteristics (e.g. PLD, spawning-mode) are similar. Schooling species of reef fish for instance were shown to produce highly variable aggregate recruitment pulses compared to species that do exhibit such adult behaviours (Shulman 1985).

**Life-history**

There are two main modes by which corals and reef fishes spawn: broadcast spawning of pelagic eggs, and spawning of benthic eggs that are brooded. Species with pelagic spawned eggs are hypothesized to have a greater potential for dispersal at least during the egg stage, a notion that is supported by the fact that fish larvae of species with pelagic eggs are typically distributed further offshore than larvae hatched from benthic eggs (Cowen & Sponaugle 1997). A recent meta-analysis of over one hundred population genetic studies on marine fish revealed significant correlations of egg type with genetic structure (Riginos et al. 2011). However, while breeding mode may be a good predictor of connectivity on evolutionary scales (F<sub>ST</sub>), it seems to exert less influence on demographic connectivity. Evidence for this stems from empirical studies in Papua New Guinea, which found little difference in the amount of self-recruitment between a benthic brooding (*Amphiprion percula*) and a pelagic spawning species (*Chaetodon vagabundus*) on a scope of 10s km (Almany et al. 2007; Berumen et al. 2012).
Likewise, genetic data from corals suggests limited correlation between reproductive mode and dispersal distances (Ayre & Hughes 2000; Miller & Ayre 2008). Further phylogeographic studies as well as more direct empirical comparisons of dispersal patterns could shed some light into the relative constraints on dispersal of pelagic versus benthic spawning species. Additionally, intensive ichthyoplankton surveys could yield fine scale information on the distribution of larvae that could be put into the context of dispersal and connectivity (Paris & Cowen 2004).

**Reef-level traits**

*Local-scale hydrodynamics*

Compared to meso-scale processes, the influence of fine-scale hydrodynamic features on connectivity is still poorly understood (Pineda et al. 2007). The already highly complex patterns of current flow in the open ocean become even more convoluted where physical processes like tides, internal waves, turbulence or wind-driven transport interact with the complex topography of a coral reef (Gawarkiewicz et al. 2007). There are few studies focusing on the near-shore hydrodynamics of reefs (e.g. Black 1993; reviewed in Monismith 2007) and fewer that combine this work with biological data (e.g. Wolanski et al. 1997).

The residence time of water around a reef likely exerts a strong influence on the dispersal of larvae from that reef and ultimately the spatial scale of connectivity (Sponaugle et al. 2002; Cetina-Heredia & Connolly 2011). Water residence is influenced by both oceanographic and topographic features, where reef-lagoons, tidal flows (Black
1993), and eddys (Leis 1986, Milicich 1994) can influence the dispersal potential of larvae from a given reef (Cetina-Heredia & Connolly 2011), and in concert with behavioural patterns can act to entrain larvae from a pool of potential source reefs (Cowen et al. 2003; Paris & Cowen 2004). The potential for self-recruitment may increase in densely aggregated reef matrices, where water flow is reduced due to the ‘sticky water’ effect (Andutta et al. 2012). These kinds of data may be more important in determining the relative level of self-recruitment a reef might expect rather than determining the proportion of settlers that are not locally produced.

**Habitat**

It is well established that larvae of reef fishes exhibit specific preferences for settlement habitats (e.g., Victor 1986; Öhman et al. 1998; Pratchett et al. 2008). Habitat preferences seem to provide enhanced fitness (Booth & Wellington 1998) and habitat availability may be a good indicator for spatial diversity patterns (Holbrook et al. 2000). Sensory capabilities have conclusively been shown to enable larvae to navigate towards favourable habitat cues (reviewed in Leis et al. 2011b), but the spatial scale of such navigation has yet to be determined. Sale et al. (2005) found little correlation of habitat differences among sampling locations and recruitment variation among samples, indicating limited active settlement choice at least on a larger scope.

On larger scales, the presence of nursery habitats, such as seagrass beds and mangrove forests has been shown to increase the variety and biomass of fishes on neighbouring reefs (Mumby et al. 2004; Nagelkerken et al. 2012). The amount and variety of particular
habitat types that are present on a given reef will surely affect that reef’s “receptivity” to larvae supplied by other source reefs, although at this point there is limited direct evidence to test this hypothesis.

**Network-level traits**

*Size and spacing of reefs*

The dispersal potential of any given reef organism depends on a variety of intrinsic and external factors (some listed here). Realized dispersal of a species, however, will depend on dispersal potential as well as the availability of suitable habitats (Jones *et al.* 2009). Coral reef ecosystems are discontinuous environments arranged in fragmented habitat patches. These patches are often distributed irregularly and the density of habitat aggregations may vary substantially throughout the landscape. In this context, species with different dispersal abilities may adapt to specific habitat types. Species with low dispersal abilities (e.g. short PLDs) may perform better in dense habitat aggregations, while species with higher dispersal potential may adapt to more segregated regions (Bode *et al.* 2011).

In coral reef frameworks, the realized dispersal of a species hence depends on the location and spacing of reefs in the network and the spatial arrangement of reefs will affect the patterns of connectivity at any given scale. Using a modeling approach, Pinksy *et al.* (2012) showed that habitat isolation for instance might produce high levels of self-recruitment (closed populations) even in species with high dispersal potential. These
kinds of data are crucial for the informed development of algorithms for the siting of marine protected areas (Leslie et al. 2003; Botsford et al. 2009).

Conditions in the pelagic environment

Similarly (or additionally) to parental effects on larval phenotypes (see above), conditions encountered in the dispersal environment might induce variability in larval quality (Sponaugle et al. 2006; Hamilton et al. 2008, Shima and Swearer, 2010). Species with feeding larvae in particular might be influenced by environmental conditions (Shima & Swearer 2009). Among many potential factors (e.g. predator regime, water chemistry), temperature (e.g. Sponaugle et al. 2006) and food availability (e.g. Shima & Swearer 2009) during the pelagic phase in particular have been identified to affect larval phenotypes and influencing dispersal patterns. Where larval quality is variable according to shifting external conditions, dispersal trajectories might be affected due to altered behavior and/or swimming capabilities, differential survival (e.g. predator avoidance), or variable success in locating a settlement site. Thus understanding the role of spatial and temporal variations in the pelagic environment pertaining to larval quality may be critical to accurately predict patterns of demographic connectivity (see Chapter 6).

Meso-scale hydrodynamics

Hydrodynamic processes acting at the meso-scale level are probably the most important physical forces governing connectivity among reefs because this is the scale at which most dispersal is occurring (James et al. 2002, Sponaugle et al. 2002, Cowen et al. 2006),
and it is most relevant to the scales of reef management. However, it has been shown that modeling ocean currents alone is not enough to accurately describe connectivity among reefs, and these models are much more powerful when coupled with other traits that are likely to influence dispersal and connectivity (Leis 2007; Werner et al. 2007). Corresponding to the local scale biophysical retention mechanisms mentioned above, behavioural attributes (e.g. vertical migration) might work in concert with meso- to large-scale hydrodynamics to restrict dispersal distances of species with theoretically high dispersal capacities (Butler et al. 2011).

Conclusions:

The above sections indicate nothing but a glimpse into the complexity of processes and mechanisms interacting on disparate temporal and spatial scales to produce patterns of connectivity on coral reefs. Each of the above points deserves an individual review and many more are barely touched or not mentioned at all (e.g. larval schooling); not to mention post-settlement processes, that eventually shape reproductive connectivity (sensu Pineda et al. 2007).

We are still far from understanding any of these stochastic factors entirely and the integration of all of them at once into an all-inclusive model seems far-fetched. While we know much more than we did ten years ago, the parameterization of individual-based dispersal models is still hampered by a lack of understanding of the interaction between larval behaviour and physical forcing. Most current models for instance include mortality and behaviour as constant inputs leveled across cohorts, an assumption that is as
unrealistic as it is hard to ameliorate. How can we ever empirically measure dynamic mortality rates over time and space? How representative are larval behavioural patterns assessed under idealistic laboratory conditions? Overall, coupled biophysical models are the future of connectivity work on larger temporal and spatial scales, however, these models should be integrated with field-based empirical evidence to support model predictions.
References:


Development and testing of 35 novel microsatellite markers for the two-band anemonefish *Amphiprion bicinctus*

**Abstract:** To investigate population genetic patterns of the anemonefish *Amphiprion bicinctus* in the Red Sea, we isolated and characterized 35 microsatellite loci using 454 sequencing. Microsatellite sequences were identified using the Tandem Repeats Finder program. Out of 100 putative markers, we identified the 35 most polymorphic and reliably scorable loci. The markers were tested on 80 individuals sampled from two spatially separated populations along the Saudi Arabian Red Sea coast. We found a mean of 10.9 alleles per locus and observed levels of heterozygosity ranged from 0.4 to 0.98. All loci were polymorphic, none deviated significantly from Hardy-Weinberg equilibrium, no linkage disequilibrium was observed and there was no evidence for null-alleles in both populations. As an additional test, we assessed within-group pairwise relatedness. The markers reported here constitute the first specific set for this species, and they are expected to contribute to future studies of connectivity in the Red Sea region.

**Introduction:**

The two-band anemonefish *Amphiprion bicinctus* is endemic to the Red Sea and parts of the Gulf of Aden. *A. bicinctus* lives in symbiosis with five species of sea anemone, *Entacmaea quadricolor*, *Stichodactyla haddoni*, *Heteractis magnifica*, *Heteractis aurora*, and *Heteractis crispa*, which occur in coral reef systems across the Indo-Pacific. After a pelagic larval phase of around 11 days, *A. bicinctus* larvae settle into a host anemone,
from which they rarely stray further than a few meters for the rest of their lives. Owing to this particular life history, anemonefish present ideal model organisms for research on patterns of connectivity, both on demographic and evolutionary scales (e.g. Buston et al. 2007, Pinsky et al. 2010; Berumen et al. 2012; Saenz-Agudelo et al. 2012). Here we developed microsatellite markers for *A. bicinctus* to study larval dispersal via parentage analysis and genetic population structure in the Red Sea.

Preceding analysis of a subset of the population genetic data of Chapter 4 in the laboratories of the Center de Biologie et d’Ecologie Tropicale et Méditerranéenne at the University of Perpignan, based on microsatellites originally developed for *Amphiprion polymnus*, did not yield reliable results. We thus decided to develop a new, species-specific collection of loci at the laboratories of KAUST.

**Methods:**

Genomic DNA from all samples mentioned in this paper was extracted from fin clip tissues using Qiagen DNeasy kits according to the manufacturer’s protocol. 454 GS FLX titanium shotgun sequencing was performed on one individual collected at Thala Reef (Population 2, see below). All procedures were conducted in the Bioscience Core Lab at the King Abdullah University of Science and Technology (Thuwal, Saudi Arabia). The genomic library for this analysis was constructed following manufacturer’s protocol.

Raw unassembled reads were mined for putative microsatellite loci using the Tandem Repeats Finder software (Benson 1999). We restricted our search to perfect di- and tri-nucleotide repeats that were at least 8 bp long. Subsequently, the Primer3 program
(Rozen & Skaletsky 2000) was used to design primers for all reads that contained suitable microsatellites repeat patterns. This way we designed primers for 100 putative microsatellite loci, which were tested in polymerase chain reaction (PCR) trials.

The 100 loci were initially tested for amplification and polymorphism on six samples collected from Palace Reef. PCRs were performed following standard protocols for the multiplex PCR kit (Qiagen) with annealing temperatures ranging from 55 to 63 °C. PCR products were then run on a Qiaxcel genetic analyzer (Qiagen) using a high-resolution cartridge to check for the presence of clear, unambiguous, polymorphic PCR products. Out of the 100 tested markers, 40 successfully amplified across the 6 samples and were polymorphic. For these loci, forward primers were labeled with fluorescent tags (6-FAM, PET, NED, VIC) and were tested on 80 samples collected from two populations (40 individuals each) located along the Saudi Arabian Red Sea coast (Fig. 1).

Each PCR was performed in 10 µl total volume containing 0.5 µl of genomic DNA, 5.0 µl of Qiagen Multiplex PCR Master Mix, 4.0 µl of H₂O, and 0.5 µl of primer mix (each primer at 2µM). Thermal profiles consisted of a denaturation step at 95 °C for 15 min, followed by 25 cycles of 30 s at 94 °C, annealing for 90 s at a locus-specific temperature (Table 1) and extension of 60 s at 72 °C, with final extension of 30 min at 60 °C. Fragment analysis was conducted in an Applied Biosystems 3730 XL genetic analyzer and allele sizes were scored using GeneMapper 4.0 software (Applied Biosystems).

Allelic frequencies, number of alleles (A), observed (H₀), and expected (Hₑ) heterozygocities were estimated for each population using the software Genalex v6.5
(Peakall & Smouse 2012). Levels of heterozygosity, departures from Hardy-Weinberg equilibrium (HWE), and linkage disequilibrium (LD) were analyzed in Genepop on the web v4.0 (Raymond & Rousset 1995; Rousset 2008). Scoring error due to stuttering, allele dropout, and presence of null alleles were tested using Microchecker v2.2.3 (Van Oosterhout et al. 2004).

**Figure 1:** Sampling locations: population 1 = Nuwayshiziyah (Lat: 26.624117, Long: 36.095317), population 2 = Thala (Lat: 22.274673, Long: 39.048463); geographic distance ~570 km.
<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequence (5′–3′)</th>
<th>Repeat motif</th>
<th>$T_a$ (°C)</th>
<th>Size range (bp)</th>
<th>$A$</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Genbank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abic01</td>
<td>F: CTCGGAGCAGAGGATGAAAG R: TAGATGAGCAGCGATTCAGTG</td>
<td>(TG)$_{10}$</td>
<td>57</td>
<td>267-365</td>
<td>23</td>
<td>0.850</td>
<td>0.858</td>
<td>KC193104</td>
</tr>
<tr>
<td>Abic02</td>
<td>F: CTCGGAGCAGAGGATGAAAG R: AGCCCATTTCTCAATGACGTCAG</td>
<td>(TG)$_{12}$</td>
<td>57</td>
<td>157-169</td>
<td>7</td>
<td>0.725</td>
<td>0.722</td>
<td>KC193105</td>
</tr>
<tr>
<td>Abic03</td>
<td>F: AGATGAGCAGCGATTCAGTG R: TTAAGAAGCAACCAGGAGCC</td>
<td>(TTA)$_{12}$</td>
<td>57</td>
<td>265-295</td>
<td>10</td>
<td>0.950</td>
<td>0.858</td>
<td>KC193106</td>
</tr>
<tr>
<td>Abic04</td>
<td>F: AAGCCCATTTGACACCCCTAC R: AGCTCCTGCTCAGTCTCTCG</td>
<td>(TG)$_{11}$</td>
<td>57</td>
<td>174-236</td>
<td>24</td>
<td>0.925</td>
<td>0.873</td>
<td>KC193107</td>
</tr>
<tr>
<td>Abic05</td>
<td>F: CCCCCAGCTTCAATAGCAG R: GCATTATGAACGTGTCTGCG</td>
<td>(ATT)$_{11}$</td>
<td>57</td>
<td>140-160</td>
<td>10</td>
<td>0.875</td>
<td>0.801</td>
<td>KC193108</td>
</tr>
<tr>
<td>Abic06</td>
<td>F: TCCCCTCCGACTGACTGAAC R: GACACACACTCCTGCTCACG</td>
<td>(ATT)$_{11}$</td>
<td>57</td>
<td>181-214</td>
<td>8</td>
<td>0.475</td>
<td>0.546</td>
<td>KC193109</td>
</tr>
<tr>
<td>Abic07</td>
<td>F: GCCAGATAAACACCCCACAC R: TGGTTTGTAGCGCATCACCTC</td>
<td>(CA)$_{16}$</td>
<td>57</td>
<td>147-163</td>
<td>9</td>
<td>0.775</td>
<td>0.852</td>
<td>KC193110</td>
</tr>
<tr>
<td>Abic08</td>
<td>F: TCCCAAAAGTGCTTTACGCGC R: ACGATGAGGTATCCATGACG</td>
<td>(GT)$_{14}$</td>
<td>57</td>
<td>245-305</td>
<td>25</td>
<td>0.975</td>
<td>0.887</td>
<td>KC193111</td>
</tr>
<tr>
<td>Abic10</td>
<td>F: CCGCTTTCTTTAAGTGATTCG R: GGAGGAACACCAAGAACCCTAC</td>
<td>(TTA)$_{15}$</td>
<td>57</td>
<td>278-311</td>
<td>10</td>
<td>0.775</td>
<td>0.813</td>
<td>KC193113</td>
</tr>
<tr>
<td>Abic11</td>
<td>F: TGCTTTGATGAGCTTACGC R: GTTGGGAGTGAGCATGACG</td>
<td>(GT)$_{13}$</td>
<td>60</td>
<td>151-211</td>
<td>14</td>
<td>0.725</td>
<td>0.718</td>
<td>KC193114</td>
</tr>
<tr>
<td>Abic12</td>
<td>F: CACATTGTGGATCAGCGCAG R: GATTGGCAGTCATGCTTCTG</td>
<td>(AAT)$_{15}$</td>
<td>60</td>
<td>227-299</td>
<td>28</td>
<td>0.825</td>
<td>0.892</td>
<td>KC193115</td>
</tr>
<tr>
<td>Abic13</td>
<td>F: TCCCCACTCAAACGAAAGAG R: AGTGTGTAAGTCCAGCCAG</td>
<td>(AC)$_{15}$</td>
<td>60</td>
<td>191-213</td>
<td>11</td>
<td>0.800</td>
<td>0.737</td>
<td>KC193116</td>
</tr>
<tr>
<td>Abic14</td>
<td>F: TGCAGCCTCCACATAAACAC R: GGTGTAAGGAGAACAGCGG</td>
<td>(CA)$_{14}$</td>
<td>60</td>
<td>276-426</td>
<td>28</td>
<td>0.850</td>
<td>0.871</td>
<td>KC193117</td>
</tr>
<tr>
<td>Abic15</td>
<td>F: CCGTAAAGCTGAAATCCTTG R: TGGCACACCTGTCTTCTCATG</td>
<td>(TTA)$_{18}$</td>
<td>60</td>
<td>139-208</td>
<td>16</td>
<td>0.975</td>
<td>0.895</td>
<td>KC193118</td>
</tr>
<tr>
<td>Abic16</td>
<td>F: GGCAAATGAGGTGAAAGGCC R: CGACTGTGCTTGGATGTTGTG</td>
<td>(CA)$_{14}$</td>
<td>60</td>
<td>249-411</td>
<td>43</td>
<td>0.900</td>
<td>0.945</td>
<td>KC193119</td>
</tr>
<tr>
<td>Locus</td>
<td>Primer sequence (5¹–3³)</td>
<td>Repeat motif</td>
<td>Tₐ (°C)</td>
<td>Size range (bp)</td>
<td>A</td>
<td>Population 1 Hₒ</td>
<td>Hₑ</td>
<td>Population 2 Hₒ</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------</td>
<td>--------------</td>
<td>---------</td>
<td>----------------</td>
<td>---</td>
<td>-----------------</td>
<td>----</td>
<td>-----------------</td>
</tr>
<tr>
<td>Abic17</td>
<td>F: TCTTTCAAGGTACACCTCAC&lt;br&gt;R: TGTCGCTCTTGTCAGGTC</td>
<td>(AC)₁₀</td>
<td>60</td>
<td>159-175</td>
<td>6</td>
<td>0.875</td>
<td>0.753</td>
<td>0.625</td>
</tr>
<tr>
<td>Abic18</td>
<td>F: ACAAAACAAGAAACTCGCC&lt;br&gt;R: GAGGCTCTATCTTATACGC</td>
<td>(AC)₁₃</td>
<td>60</td>
<td>259-301</td>
<td>11</td>
<td>0.700</td>
<td>0.752</td>
<td>0.775</td>
</tr>
<tr>
<td>Abic19</td>
<td>F: AGTGAAGGCAATTTTCTAGC&lt;br&gt;R: AGGAGAGCAAGGGGCTAC</td>
<td>(TG)₁₁</td>
<td>60</td>
<td>150-174</td>
<td>8</td>
<td>0.725</td>
<td>0.718</td>
<td>0.775</td>
</tr>
<tr>
<td>Abic20</td>
<td>F: TGTGCACAGTGGAAGTACG&lt;br&gt;R: AGTGAAGATCTCGGTGTCG</td>
<td>(AC)₁₅</td>
<td>60</td>
<td>241-255</td>
<td>8</td>
<td>0.775</td>
<td>0.819</td>
<td>0.800</td>
</tr>
<tr>
<td>Abic21</td>
<td>F: ACTGGTGTTGCTGTTTCCTC&lt;br&gt;R: GTCCACGTGTGTCTGTGTGC</td>
<td>(AC)₁₁</td>
<td>60</td>
<td>190-214</td>
<td>11</td>
<td>0.850</td>
<td>0.832</td>
<td>0.900</td>
</tr>
<tr>
<td>Abic22</td>
<td>F: CGATTTGCTTTCGTTTTCAGC&lt;br&gt;R: AAAGCAACGAGCATCTCTG</td>
<td>(CAG)₁₇</td>
<td>60</td>
<td>231-273</td>
<td>14</td>
<td>0.900</td>
<td>0.857</td>
<td>0.850</td>
</tr>
<tr>
<td>Abic23</td>
<td>F: TTACACTCTCCTCCACCAGG&lt;br&gt;R: CGGTAGACCTGACAGGAAGC</td>
<td>(CA)₁₄</td>
<td>60</td>
<td>162-250</td>
<td>16</td>
<td>0.750</td>
<td>0.855</td>
<td>0.850</td>
</tr>
<tr>
<td>Abic24</td>
<td>F: TTTGACTCTCCTGTGTTTGACAG&lt;br&gt;R: TCCCTCTTCTGAGCTTTCG</td>
<td>(AC)₁₃</td>
<td>60</td>
<td>264-282</td>
<td>10</td>
<td>0.600</td>
<td>0.743</td>
<td>0.650</td>
</tr>
<tr>
<td>Abic25</td>
<td>F: ACAGTACCAGCTGCTACATCAAC&lt;br&gt;R: AGCAGAGCTCTTTACGCTG</td>
<td>(CA)₁₂</td>
<td>63</td>
<td>253-281</td>
<td>10</td>
<td>0.400</td>
<td>0.354</td>
<td>0.400</td>
</tr>
<tr>
<td>Abic26</td>
<td>F: TGCCAGCAAACCTCTCTACAC&lt;br&gt;R: AAAGCAACGACAGTGTCACG</td>
<td>(AGG)₁₀</td>
<td>63</td>
<td>142-163</td>
<td>6</td>
<td>0.600</td>
<td>0.618</td>
<td>0.575</td>
</tr>
<tr>
<td>Abic27</td>
<td>F: GGGGTAGCACAGAGAGACAG&lt;br&gt;R: TCCCTACGAATAGCCAGCTG</td>
<td>(CA)₁₃</td>
<td>63</td>
<td>250-272</td>
<td>11</td>
<td>0.800</td>
<td>0.740</td>
<td>0.775</td>
</tr>
<tr>
<td>Abic28</td>
<td>F: ATAGATGAGGAGACAGGAGC&lt;br&gt;R: GGCACTCCATTGGATTCTTG</td>
<td>(AAT)₁₃</td>
<td>63</td>
<td>177-222</td>
<td>12</td>
<td>0.825</td>
<td>0.845</td>
<td>0.775</td>
</tr>
<tr>
<td>Abic29</td>
<td>F: ACAGCCTGCAGCTTTATGACC&lt;br&gt;R: ATCCGACACAGAAGACCTTG</td>
<td>(AC)₁₃</td>
<td>63</td>
<td>277-287</td>
<td>6</td>
<td>0.475</td>
<td>0.580</td>
<td>0.700</td>
</tr>
<tr>
<td>Abic30</td>
<td>F: GCAGAAATGGATGTTGTTG&lt;br&gt;R: GTGGGTGTGAGTTCTGTTG</td>
<td>(TA)₁₁</td>
<td>63</td>
<td>159-167</td>
<td>5</td>
<td>0.550</td>
<td>0.717</td>
<td>0.625</td>
</tr>
<tr>
<td>Abic31</td>
<td>F: ACTCCCAACCCGAGAAGAGG&lt;br&gt;R: CTCCCTGACTTTTTAGC</td>
<td>(CA)₁₃</td>
<td>63</td>
<td>254-270</td>
<td>9</td>
<td>0.900</td>
<td>0.812</td>
<td>0.650</td>
</tr>
<tr>
<td>Locus</td>
<td>Primer sequence (5’–3’)</td>
<td>Repeat motif</td>
<td>$T_a$ (°C)</td>
<td>Size range (bp)</td>
<td>$A$</td>
<td>Population 1</td>
<td>Population 2</td>
<td>Genbank accession no.</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>---------</td>
<td>----------------</td>
<td>-----</td>
<td>---------------</td>
<td>---------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| Abic32 | F: TCCAGGATAGTGCTGCTGAG  
R: ATGGCCTTTGCTTTAACGTGG  
(CTG)$_{10}$ | 63 | 157-172 | 6 | 0.675 | 0.662 | 0.800 | 0.732 | KC193135 |
| Abic33 | F: GACACACACTGGGCTGACTG  
R: TAAAGCAGGTGCTGTGATG  
(GT)$_{11}$ | 63 | 179-209 | 11 | 0.800 | 0.711 | 0.575 | 0.628 | KC193136 |
| Abic34 | F: GAAAGGGGTAGAGAGAGAGG  
R: GACTTGGTGTGGGTGGAATC  
(AC)$_{10}$ | 63 | 246-268 | 12 | 0.625 | 0.677 | 0.750 | 0.694 | KC193137 |
| Abic35 | F: CCATGAAAAAGCAGAACACGC  
R: CTGAAGCCTCTCAGCTCACCC  
(GTG)$_{10}$ | 63 | 142-172 | 10 | 0.825 | 0.743 | 0.750 | 0.723 | KC193138 |
| Abic36 | F: ACATCACAGAGGGCAGACGAC  
R: TGGAGCAAATGAAATGGTG  
(CA)$_{14}$ | 63 | 239-283 | 13 | 0.750 | 0.816 | 0.725 | 0.767 | KC193139 |

**Table 1:** Primer sequence, repeat motif, annealing temperature ($T_a$), total allele size range, total number of alleles ($A$), expected ($H_E$) and observed ($H_O$) heterozygosities for 35 microsatellite loci tested in 80 individuals of *A. bicinctus* from two populations ($n = 40$ each)
As an additional test for the functionality of the new markers, we assessed for both populations mean heterozygosity and mean within-group pairwise relatedness (Lynch & Ritland 1999), the latter with 10,000 permutations and 1,000 bootstraps as implemented in Genalex.

To reduce analysis costs and processing time we developed and tested different multiplex PCR mixes for (1) testing the new primers (this chapter), (2) population genetic analysis, including three markers from Perpignan (Chapter 4), and (3) parentage assignment analysis (30 loci, Chapter 5). Mixes were grouped according to annealing temperatures and contained six to nine primers each.

**Results and Discussion**

Of the 40 loci tested, 35 yielded clear peaks that could be scored unambiguously and amplified in all samples. Allelic diversity ranged between 5 and 43 (mean allelic diversity was 10.91 ± 0.699SE), observed and expected heterozygosity ranged between 0.4 and 0.98 and 0.35 and 0.95 respectively (mean observed and expected heterozygosity were 0.76 ± 0.016SE and 0.76 ± 0.014SE respectively). After corrections for multiple testing via False Discovery Rate estimation (Benjamini & Hochberg 1995), no locus sowed significant deviations from HWE expectations (p-value range: 0.0014–0.05) and no significant LD was observed in both populations for any pair of loci (p-value range: 0.00008–0.05). Results from Microchecker show no evidence of null-alleles in both populations.
Mean heterozygosity did not differ significantly between Nuwayshiziyah and Thala (t-test: $H_O: p = 0.76; H_E: p = 0.42$) (Fig. 2a). Mean within-group pairwise relatedness was low for both populations ($r = -0.011$) and not significantly different (t-test: $p = 0.97$) (Fig. 2b).

The markers presented here will be useful in the study of connectivity patterns in the Red Sea, an area that has historically been understudied in terms of marine ecological research (Spaet et al. 2012).
References


Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea

Abstract: There is accumulating evidence for the effect of environmental heterogeneity on patterns of connectivity in marine systems. The emerging field of seascape genetics seeks to identify the influence of biophysical habitat features on the spatial genetic structure of populations or individuals. Here we investigate the population genetic patterns of an anemonefish, *Amphiprion bicinctus*, along the Saudi Arabian coast of the Red Sea. We collected nearly one thousand samples from 19 locations, spanning approximately 1500 km, and genotyped them at 38 microsatellite loci. Patterns of gene flow followed a stepping-stone model along the northern and central Red Sea, but were disrupted by a distinct genetic break at a latitude of approximately 19°N. The Red Sea is characterized by pronounced environmental gradients along its axis, roughly separating the northern and central from the southern basin. Using mean chlorophyll-a concentrations as a proxy for this gradient, we ran tests of isolation by distance (IBD, $R^2 = 0.52$) and isolation by environment (IBE, $R^2 = 0.64$). Moreover, we implemented partial Mantel tests and Multiple Matrix Regression with Randomization (MMRR). We found that genetic structure across our sampling sites was best explained by a combined model of IBD and IBE ($R^2 = 0.71$ and $0.92$ respectively). We support growing calls for the integration of habitat characteristics into future population genetic studies and into the parameter set for biophysical models of larval dispersal.
Introduction:

One of the fundamental goals of population genetic studies is to understand the nature of genetic structuring in real populations and ultimately identify the processes responsible for shaping those patterns. A comprehensive understanding of the factors that affect population connectivity and persistence is a key element in effective conservation management (Botsford et al. 2009; Gaines et al. 2010). The relatively recent field of seascape genetics seeks to assess the influence of environmental factors on spatial genetic divergence in marine organisms (Selkoe et al. 2008; Liggins et al. 2013). A common approach to seascape genetics is the model of isolation by distance (IBD), whereby genetic differentiation is expected to increase linearly with geographic distance in continuous populations under an equilibrium between migration and drift (Wright 1943; Slatkin 1993). IBD may, to some extent, reconcile the gap between direct (e.g. parentage analysis) and indirect genetic approaches by estimating ecologically relevant parameters of larval dispersal (Rousset 1997; Pinsky et al. 2010). The model is indeed applied regularly in marine systems with varying degrees of genetic variability explained by geographic distance (Bradbury & Bentzen 2007; Selkoe & Toonen 2011), yet several studies have found no spatial-genetic relationship at all (e.g. Doherty et al. 1995; Selkoe et al. 2006; Bradbury & Bentzen 2007; Selkoe et al. 2010). The lack of any obvious genetic pattern in space is termed ‘chaotic genetic patchiness’ (Johnson & Black 1982) and indicates the importance of other factors besides spatial distance in shaping genetic structure (Selkoe et al. 2010; Liggins et al. 2013). Certain environmental processes and habitat characteristics may thus influence gene flow either directly by disrupting dispersal
(e.g. physical barriers), or indirectly by preventing successful settlement or subsequent survival (e.g. local adaptation) (Nosil et al. 2005; White et al. 2010; Burgess et al. 2012).

Undoubtedly, oceanic currents are key drivers in the formation, maintenance, or disruption of genetic patterns in marine systems (e.g. Galindo et al. 2006; White et al. 2010), yet other less investigated factors may play equally important roles (Riginos & Liggins 2013). Isolation by environment (IBE, Wang & Summers 2010) is increasingly becoming the focus of seascape genetic studies, correlating environmental parameters with genetic divergence (e.g. Alberto et al. 2010; Mendez et al. 2010). IBE predicts a positive relationship between genetic and environmental dissimilarity, mainly because greater habitat differentiation among populations is expected to reduce the fitness of dispersers, thereby leading to stronger divergent selection (Marshall et al. 2010; Bonte et al. 2012; Wang et al. 2013). Accumulating evidence for shifts in allele frequencies along environmental gradients in marine systems (reviewed in Schmidt et al. 2008) are indicative of the influence of selective processes in shaping population genetic patterns under an IBE model (Riginos & Liggins 2013). Inherently, IBD and IBE are not mutually exclusive concepts and spatially structured population divergence may be a result of a combination of geography and environment (Crispo et al. 2006; Alberto et al. 2010; Wang et al. 2013). A better comprehension of the synergistic effects of IBD and IBE on patterns of gene flow among populations will inform our understanding of the effects of environmental heterogeneity on patterns of genetic variation in nature (Wang & Summers 2010).

The Red Sea has long been recognized for its extensive, diverse coral reef systems and its high levels of endemism (DiBattista et al. 2013). Nevertheless, the region remains
remarkably understudied in terms of marine ecological research (Spaet et al. 2012; Berumen et al. 2013). The Red Sea basin is characterized by pronounced biophysical gradients along a nearly linear coastline (Raitsos et al. 2013), presenting an excellent opportunity for the investigation of the effects of geographic distance and environmental heterogeneity on marine population connectivity. At approximately south of 20°N, the Red Sea exhibits a marked increase in turbidity accompanied by a pronounced decrease in reef development. Previous surveys have found swift changes in the structure of reef fish assemblages (Roberts et al. 1992) as well as a genetic break in a common reef fish species (Froukh & Kochzius 2007) that coincide markedly with this environmental shift. Yet so far it has not been possible to statistically infer a pattern of IBE.

The application of approaches derived from terrestrial landscape genetics in the ocean remains challenging, mainly due to the dynamic nature of oceanography and the often delicate genetic structure resulting from large effective population sizes and extensive connectivity (Hellberg et al. 2002; Selkoe et al. 2008). Increasing the number of variable genetic markers or the amount of individuals sampled per population were both shown to increase the power of landscape genetic inferences by detecting subtle genetic structure in situations of high gene flow (Waples 1998; Landguth et al. 2012).

Here we investigated the genetic structure of a common reef fish, the two-band anemonefish *Amphiprion bicinctus*, with a relatively short pelagic larval duration (PLD = 12 days) along a nearly continuous reef system. We used the highest number of highly variable, polymorphic microsatellites (38 loci) in a coral reef fish genetic survey to date in concert with an extensive sampling regime on a scale of hundreds of kilometers along the Saudi Arabian Red Sea coast. By doing this, we were able to test and tease apart the
individual and combined effects of IBD and IBE on patterns of genetic differentiation in a marine ecosystem.

**Materials and Methods:***

*Study system*

The Red Sea is a narrow ocean basin, embedded between the African and Asian continental shelves that extends approximately 2250 km between 30° and 13° latitude. It connects to the Mediterranean Sea through the Suez Canal in the north and to the Gulf of Aden and the Indian Ocean through the strait of Bab-el-Mandeb in the south. A deep trench, the Red Sea Rift, stretches along its entire north-south axis, resulting in depths in excess of 2200 m. Due to its hot and arid climate, the Red Sea experiences very little freshwater inflow and high rates of evaporation, rendering it one of the warmest and saltiest ocean basins in the world (Johns *et al.* 1999).

The Red Sea further exhibits pronounced latitudinal and seasonal gradients in physical and environmental parameters. Some examples include north-south gradients in sea surface temperature (summer: 26-32°C; winter: 20-28°C), salinity (42-37‰), primary productivity and turbidity (low-high) (Raitsos *et al.* 2011, 2013). The basin is thus characterized by a distinct habitat shift marked by steep coasts and clear, oligotrophic waters in the northern and central Red Sea and shallow banks with turbid, nutrient-rich waters in the south (Appendix A).
Figure 1: Sampling locations along the Saudi Arabian coast of the Red Sea. Refer to Table 1 for details about the sites. Division of the basin into northern, central and southern Red Sea according to a separated IBD model (see Fig. 5).

Here we considered the two-band anemonefish, *Amphiprion bicinctus* (Rüppel 1830), a common inhabitant of coral reefs in the Red Sea, which occurs in association with five species of anemone, *Entacmaea quadricolor, Stichodactyla haddoni, Heteractis magnifica, Heteractis aurora, and Heteractis crispa* between 0.5 and 45 m depth (Fautin and Allen 1992). While its host anemones occur throughout the Indo-Pacific region, *A. bicinctus* is endemic to the Red Sea and parts of the Gulf of Aden. The Saudi Arabian Red Sea coast presents a nearly continuous habitat for *A. bicinctus* with fringing reefs along most of its entire length.
Sampling

A total of 991 samples were collected between October 2009 and August 2011 from 19 locations spanning more than 1500 km along the Saudi Arabia coast of the Red Sea (Fig. 1; Table 1). Individual specimens encountered during SCUBA dives were caught using aquarium hand nets and clove oil, clipped at the caudal fin using conventional scissors and released back into their host anemone. Fin clips were stored in 96% ethanol and transported back to the laboratory for subsequent genetic analysis.

Genetic analysis

All samples were genotyped over 38 microsatellite loci, 35 of which were developed by Nanninga et al. (2012). The additional three unpublished loci had originally been developed for *Amphiprion polymnus* at the Center de Biologie et d’Ecologie Tropicale et Méditerranéenne at the University of Perpignan (Table 2). Genomic DNA was extracted and fragments amplified, sequenced and scored according to Nanninga et al. (2012).

The data was tested for the presence of null-alleles using Microchecker (v2.2.3; van Oosterhout et al. 2004). Deviations from Hardy-Weinberg-Equilibrium (HWE) and Linkage Disequilibrium (LD) between loci were tested in GENEPOP (web-service; Raymond & Rousset 1995), adjusting for multiple tests using False Discovery Rates (FDR, Benjamini & Hochberg 1995). Expected, observed and unbiased heterozygosities as well as allelic richness were calculated using Genalex (v6.5; Peakall & Smouse 2012).

We conducted an analysis of molecular variance (AMOVA) to investigate population structure as implemented in Genalex. Genetic structure was evaluated among regions
(North, Central, South; Fig. 1) and among populations within regions. Both global and pairwise $F_{ST}$ were tested for significance with 10,000 permutations.

**Table 1:** Sampling regime with original site names, abbreviation codes, GPS coordinates, number of samples collected per site and date of collections.

<table>
<thead>
<tr>
<th>Site name</th>
<th>Code</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Samples</th>
<th>Date sampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burcan</td>
<td>Bur</td>
<td>27°54'35.46&quot;</td>
<td>35°03'55.20&quot;</td>
<td>58</td>
<td>August 2011</td>
</tr>
<tr>
<td>An Numan</td>
<td>AN</td>
<td>27°08'11.64&quot;</td>
<td>35°45'03.06&quot;</td>
<td>56</td>
<td>August 2011</td>
</tr>
<tr>
<td>Nuwayshiziyah</td>
<td>Nuw</td>
<td>26°37'26.82&quot;</td>
<td>36°05'43.14&quot;</td>
<td>53</td>
<td>August 2011</td>
</tr>
<tr>
<td>Mashabi</td>
<td>Mas</td>
<td>25°34'56.70&quot;</td>
<td>36°32'55.08&quot;</td>
<td>54</td>
<td>August 2011</td>
</tr>
<tr>
<td>Abu Matari</td>
<td>Amat</td>
<td>24°43'23.82&quot;</td>
<td>37°09'04.20&quot;</td>
<td>53</td>
<td>August 2011</td>
</tr>
<tr>
<td>Yanbu</td>
<td>Yan</td>
<td>24°08'58.29&quot;</td>
<td>37°40'30.18&quot;</td>
<td>51</td>
<td>October 2010</td>
</tr>
<tr>
<td>Shi’b al Bayda</td>
<td>SB</td>
<td>22°44'32.96&quot;</td>
<td>38°46'57.42&quot;</td>
<td>60</td>
<td>January 2010</td>
</tr>
<tr>
<td>Haitham</td>
<td>Hai</td>
<td>22°16'28.82&quot;</td>
<td>39°02'54.47&quot;</td>
<td>58</td>
<td>January 2010</td>
</tr>
<tr>
<td>Abu Madafi</td>
<td>AM</td>
<td>22°04'10.64&quot;</td>
<td>38°46'18.04&quot;</td>
<td>60</td>
<td>January 2010</td>
</tr>
<tr>
<td>Palace Reef</td>
<td>PR</td>
<td>22°13'25.53&quot;</td>
<td>38°58'07.67&quot;</td>
<td>46</td>
<td>January 2010</td>
</tr>
<tr>
<td>Eagle Reef</td>
<td>ER</td>
<td>21°48'52.72&quot;</td>
<td>38°50'15.14&quot;</td>
<td>56</td>
<td>January 2010</td>
</tr>
<tr>
<td>Abu Terr</td>
<td>AT</td>
<td>21°40'35.24&quot;</td>
<td>38°50'28.76&quot;</td>
<td>59</td>
<td>January 2010</td>
</tr>
<tr>
<td>Tawil Ral</td>
<td>TR</td>
<td>20°38'51.38&quot;</td>
<td>39°23'41.42&quot;</td>
<td>44</td>
<td>October 2009</td>
</tr>
<tr>
<td>Um Haj</td>
<td>UH</td>
<td>20°22'07.48&quot;</td>
<td>39°39'02.13&quot;</td>
<td>59</td>
<td>October 2009</td>
</tr>
<tr>
<td>Canyon Reef</td>
<td>Can</td>
<td>19°53'25.32&quot;</td>
<td>39°57'38.40&quot;</td>
<td>49</td>
<td>October 2009</td>
</tr>
<tr>
<td>AQ3</td>
<td>AQ</td>
<td>19°06'31.92&quot;</td>
<td>40°29'20.94&quot;</td>
<td>55</td>
<td>October 2009</td>
</tr>
<tr>
<td>Maghabiyah</td>
<td>Mag</td>
<td>18°16'22.20&quot;</td>
<td>40°44'11.22&quot;</td>
<td>32</td>
<td>October 2009</td>
</tr>
<tr>
<td>Sumayr</td>
<td>Sum</td>
<td>17°47'14.46&quot;</td>
<td>41°26'30.60&quot;</td>
<td>32</td>
<td>October 2009</td>
</tr>
<tr>
<td>Farasan</td>
<td>Far</td>
<td>16°37'05.41&quot;</td>
<td>41°56'16.26&quot;</td>
<td>56</td>
<td>November 2010</td>
</tr>
</tbody>
</table>

STRUCTURE (v2.3.4; Pritchard et al. 2000) was used to estimate the number of genetically differentiated clusters ($K$). The algorithm assigns all individuals in the total sample to clusters that are estimated using a log-likelihood approach of pre-defined $K$ values to minimize HW disequilibrium. We used an admixture model with sampling site as location prior. Allele frequencies were assumed to be correlated among populations. Default values for alpha and prior $F_{ST}$ were used. Log-likelihood values were computed for each $K$ (1-10) by running STRUCTURE 10 times with 200,000 repetitions each.
STRUCTURE Harvester (Earl 2012) was used to visually assess the number of clusters $K$ that best describes our data according to the highest averaged maximum log-likelihood.

Table 2: Primer sequence, repeat motif, annealing temperature ($T_a$), total allele size range, total number of alleles ($A$), expected ($H_e$) and observed ($H_o$) heterozygosities for three previously unpublished microsatellite loci ($n = 991$).

<table>
<thead>
<tr>
<th>Loci</th>
<th>Primer sequence (5’–3’)</th>
<th>Repeat motif</th>
<th>$T_a$ (˚C)</th>
<th>Size range (bp)</th>
<th>$A$</th>
<th>$H_o$</th>
<th>$H_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4</td>
<td>F: TTGTTACTGTGTCCGTGTGATC R: GGCGACATGATACACTTGACTT</td>
<td>(GT)$_{15}$</td>
<td>60</td>
<td>170-194</td>
<td>11</td>
<td>0.85</td>
<td>0.86</td>
</tr>
<tr>
<td>A108</td>
<td>F: TTGCTTCGAGTTCATAGAC R: CACAACACACACACACATTGTA</td>
<td>(TG)$_{12}$</td>
<td>57</td>
<td>287-329</td>
<td>19</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>A109</td>
<td>F: CAGCATTAGTGCCATTGTC R: TATAGGCGAGAATGAAGCGAAC</td>
<td>(CA)$_{15}$</td>
<td>57</td>
<td>251-283</td>
<td>15</td>
<td>0.95</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Spatial genetic structure has traditionally been described by summary statistics between pairs of populations. A more recent approach is centered on a network perspective, allowing for the analysis of the statistical relations among all populations simultaneously (Dyer & Nason 2004). Network analysis was implemented in EDENetworks (v2.16; M. Kivelä, S. Arnaud-Haond, J. Saramäki, in preparation) to construct a minimum spanning tree (MST). An MST is the minimal network necessary to connect all populations in the sample. For this purpose, the program plots all populations (nodes) in a network graph with connections (edges) between all nodes. Each edge is weighted according to its pairwise genetic distance ($F_{ST}$). The MST selects a subset of edges that connects all nodes while minimizing the overall genetic distance. By analyzing the shape of the MST we can make inferences about the nature of gene flow through the study system. Network construction was based on genetic distance between populations.
with no further priors and the layout of the MST recalculated 10 times to test for possible alternative network shapes.

**Isolation by distance**

We tested for isolation-by-distance (IBD) by implementing Mantel tests based on difference matrices of pairwise linearized $F_{ST} (F_{ST} / (1 - F_{ST}))$ and Euclidean distance (km) between all sampling locations in Genalex with 10,000 permutations.

To test whether IBD is the main force shaping patterns of genetic differentiation among our sampling sites along the Saudi Arabian coast, we divided the Red Sea into three regions, North, Central and South (Fig. 1). Within each region, we plotted IBD patterns as above and ran a generalized linear model (GLM) with “region” as covariate. If we assume consistent sampling as well as a uniform coastline across sections we would expect the slopes of the functions to be homogenous among regions if IBD was the key driver of genetic structure along the Red Sea. Varying slopes on the other hand might indicate that other forces besides IBD might influence the observed genetic patterns.

**Isolation by environment**

As a representative parameter describing the marked gradient in habitat characteristics between the southern and the northern and central Red Sea, we chose chlorophyll a concentrations (chl-a). While there are several physical parameters that constitute the environmental gradient in the Red Sea, chl-a coincides best with the observed habitat change. Salinity, for instance, changes gradually along the axis of the Red Sea and hence correlates strongly with geographic distance.
Validated mean measures of chl-a (mg/m$^3$) at all sampling locations were obtained from D. Raitsos (pers. comm.). Based on a 10-year high-resolution data set of satellite remote sensing chl-a estimates, Raitsos et al. (2013) provided a detailed description of the patterns of chl-a succession in the Red Sea. For statistical analysis we calculated a pairwise distance matrix of chl-a concentrations between sampling locations.

There is an ongoing debate about the best approach towards the analysis of non-independent, pairwise data (Legendre & Fortin 2010). Locations or populations that are geographically more distant inherently tend to experience larger differences in environmental parameters. Disentangling the effects of IBD and IBE is hence challenging (Wang et al. 2013). Partial Mantel tests relate two main matrices while controlling for the effects of another (Smouse et al. 1986). While this approach makes few assumptions and is robust towards the potential additive effects of related distance matrices, it lacks the statistical power of regression and correlation approaches (Legendre & Fortin 2010). Here we implemented Multiple Matrix Regression with Randomization (MMRR) according to Wang (2013). MMRR is conceptually similar to partial Mantel tests but incorporates multiple regression, thereby allowing for the statistical quantification of the additive effects of multiple predictors (Wang 2013).

To assess the effect of geographic and environmental distance (chl-a) on genetic variability, we first implemented partial Mantel tests in IBDWS (web service; Jensen et al. 2005) with 10,000 permutations. We then implemented MMRR in R (available from the Dryad Data Repository; doi:10.5061/dryad.kt71r) to assess the additive effect of both independent factors (geographic distance and chl-a). All matrices were standardized using the “scale” function implemented in R before running the MMRR.
Results:

General genetic patterns

Genetic diversity within populations of *A. bicinctus* was generally high (mean $H_o = 0.755$, mean $H_E = 0.754$) with little variation among the northern and central locations, but with a steep drop in heterozygosity south of 19° latitude (Fig. 2). The number of alleles ranged from 4 to 44 and patterns of mean allelic richness closely mirrored patterns of heterozygosity with a drop from 12 to 10 at 19°N.

Tests for linkage disequilibrium (LD) were significant in 61 of 13,357 comparisons (0.005%) after correcting for multiple tests (FDR). There were no locus specific patterns across populations. There were significant deviations from HWE in 3 of 722 comparisons (0.004%) after FDR. No locus deviated significantly from HWE in more than two populations suggesting the absence of null alleles. All loci were included for subsequent analyses.

Despite a low global $F_{ST}$ (0.004), AMOVA revealed significant genetic structure in the data ($p = 0.001$). Inclusion of region in the analysis increased the global $F_{ST}$ to 0.007 ($p = 0.001$). The source of genetic variability was 99.35% within populations, 0.21% among populations, and 0.44% among regions (North, Central, South). Pairwise comparisons among locations revealed little structure among the northern and central populations (mean $F_{ST} = 0.0019$), but half an order of magnitude higher differentiation between the southern and the northern and central populations (mean $F_{ST} = 0.0111$) (Table 3). Out of all 19 locations, Farasan was the most differentiated (see Appendix B for alternative measures of genetic distance – $F'_{ST}$ and Nei’s distance).
Table 3: AMOVA table of pairwise $F_{ST}$ (below diagonal) and significance values ($p$) after 10,000 permutations (above diagonal). Significant values are bold.
Analysis in STRUCTURE revealed the first and only local maximum of $P(K)$ at $K = 2$ (Appendix C). Assignments of individuals according to $K = 2$ clusters show a steady gradient from one cluster to the other from north to south, with a clear leap from AQ to Maghabiyah and again a more pronounced one from Sumayr to Farasan (Fig. 3). This is in conformity with AMOVA results of high differentiation between the southern most populations and the rest.

![Figure 2: Mean unbiased heterozygosities (uHe ±1SE) against latitude across sampling locations from north to south.](image)

Network analysis revealed a minimum spanning tree that conforms notably to the geographic setting of the sampled locations in the absence of any prior spatial input parameters (Fig. 3). Repeated recalculation of the layout did not alter the overall shape of the tree. The conformity between geographic and MST layout indicates stepping-stone gene flow among our sampling sites. The tree shows the highest amount of ‘connectedness’ among the central populations with decreasing values towards the edges.
of the system. As in previous analyses, the three southernmost populations appear less connected than the rest.

**Figure 3**: Minimum Spanning Tree (MST) (a) and STRUCTURE bar plot based on $K = 2$ (b). While the MST closely resembled the layout of the sampling locations, the bar plot is aligned from north to south for ease interpretation. MST: size of nodes and edges are scaled to the degree of “connectedness” to other populations.

*Isolation by geographic and environmental distance*

Division of the coastline into regions revealed variable patterns of IBD between the northern and central and the southern Red Sea (Fig. 4). GLM results indicate that for the northern and central regions, IBD slopes were significantly different from zero ($p = 0.018$), but not different from each other (North: $m = 8 \times 10^{-6}$, Central: $m = 6 \times 10^{-6}$, $p = 0.704$). In contrast, the correlation slope increased by an order of magnitude in the south (South: $m = 5 \times 10^{-5}$, $p < 0.0001$) (Fig. 4). These notably different IBD slopes across a
relatively homogenous coastline point towards additional factors that influence genetic differentiation along the Saudi Arabian Red Sea coast besides geographic distance.

Mantel tests of IBD over all populations revealed a significant ($p < 0.0001$), positive ($m = 0.00001$) relationship between geographic and linearized genetic distances. The model explained over 50% of the variation in the data ($R^2 = 0.522, p < 0.0001$) (Fig. 4). Chl-a was an even better predictor than geographic distance with Mantel tests explaining 64% of genetic variability along the Red Sea coast ($R^2 = 0.64, p < 0.0001$). The model was improved by integrating both predictors in a partial Mantel test ($R^2 = 0.71, p < 0.0001$).

**Figure 4:** Functions of IBD: a) across the entire data set, b) within three regions (see Fig. 1) of the Red Sea (blue squares: North, red dots: Central, green triangles: South).
MMRR revealed that IBE ($\beta_E = 0.68, p = 0.001$) was more important in explaining the genetic pattern than IBD ($\beta_D = 0.44, p = 0.001$). Both factors together explained over 90% of the genetic variability ($R^2 = 0.918, p = 0.001$) (Fig. 5). We found a moderate but significant correlation between geographic distance and chl-a ($R^2 = 0.127, p = 0.014$).

**Figure 5:** MMRR plots of (a) the relationship between geographic and environmental distance, (b) IBD, (c) IBE and (d) the combined effects of geographic and environmental distance on genetic distance.
Discussion:

Our data and combined analyses provide explicit evidence that environmental gradients may influence patterns of gene flow in coral reef fishes with pelagic larval dispersal. Employing the highest number of polymorphic microsatellite markers used to date in a coral reef fish population genetic study we detected subtle genetic patterns, which were congruent among a number of statistical approaches.

The most prominent observed pattern was high genetic connectivity among the northern and central locations with a significant ($p = 0.001$) genetic break at around 19° latitude (Fig. 2, 3; Table 3). It should be noted that the magnitude of this break is low (highest pairwise $F_{ST} = 0.031$), suggesting that there is a certain level of gene flow between the southern and northern populations.

Genetic structure between the northern and southern Red Sea has been identified before. Froukh and Kochzius (2007) sampled fourline wrasse, *Larabicus quadrilineatus*, at five locations along the Saudi Arabian coast. They identified a genetic break between the northern and the central and southern Red Sea at around 20°N. The discrepancy of one degree latitude between the studies might be due to the lower spatial sampling resolution as well as potentially species-specific dispersal patterns or tolerances to environmental gradients.

Despite low levels of genetic differentiation overall, this apparent discontinuity to gene flow was identified both via AMOVA and Bayesian cluster analysis (STRUCTURE, Fig. 3). The three southernmost populations markedly differed from the other populations in terms of genetic diversity and genetic differentiation, with the
southernmost population Farasan particularly standing out (Fig. 3, Table 3). This apparent break in gene flow coincides with a notable environmental shift along the axis of the Red Sea (Raitsos et al. 2013; Appendix A). Using chl-a concentrations as a proxy for this habitat change, we showed that population structure in our system is best explained by the synergistic effects of geographic and environmental distances on patterns of gene flow (Fig. 5).

Interestingly, the slope of IBD was significantly higher in the southern region than in the northern and central regions (Fig. 4). These differences in regional patterns suggest that IBD is most likely not the only factor shaping genetic structure along the Saudi Arabian coast, even though geographic distance alone explained over 50% of the variation in our data over all populations ($R^2 = 0.522$; Fig. 4). Despite this already high level of model fit with only one predictor, the model was improved significantly by incorporating differences in average chl-a concentrations among sampling locations (Mantel: $R^2 = 0.71$, MMRR: $R^2 = 0.92$; Fig. 5).

Chl-a concentrations were shown to be a good predictor of larval fish assemblage composition (Carassou et al. 2008), yet there is little information on the influence of primary productivity on settlement success. Lo-Yat et al. (2011), observed a positive relationship between chl-a and larval supply, probably related to the higher food abundance in high productivity environments. On the other hand, changes in turbidity have been shown to impair habitat choice and foraging success of coral reef fish larvae by impairing the detectability of visual cues (Wenger et al. 2011, 2012).
One needs to be cautious when interpreting correlations among distance matrices and we do not infer a direct causal influence of chl-a on the genetic patterns of *A. bicintus* in the Red Sea. Chl-a rather serves as a representation of the pronounced environmental gradient from clear, oligotrophic, steep reefs in the northern and central Red Sea to a turbid, nutrient rich, shallow shelf in the south (Roberts *et al.* 1992; Raitsos *et al.* 2011, 2013). While this habitat shift might not act as a physical barrier against migration, it might disrupt successful settlement or survival of immigrants (Nosil *et al.* 2005; Marshall & Morgan 2011).

Genetic diversity (unbiased heterozygosity) decreased markedly in the southern populations (Fig. 2), indicating a comparatively low level of gene flow into this area and may be a sign of selective sweepstakes. This hypothesis goes in line with previous findings of northern gene flow directionality along the southern Saudi Arabian coast (Froukh & Kochzius 2007). Network analysis additionally underlined those patterns by indicating markedly decreasing connectedness towards the southern populations (Fig. 3). This relative dearth of successful immigrants might result from selection against recruits from non-matching natal environments, arising from a competitive disadvantage of genotypes adapted to their natal habitat (Nosil *et al.* 2005; Vigliola *et al.* 2007; Marshall *et al.* 2010).

While our model captures a substantial part of the variability in the data, we do not dismiss the possible synergistic or antagonistic influence of oceanography. Indeed, the genetic break between the southern and central and northern populations roughly coincides with two previously described semi-persistent eddies (cyclic and anticyclic), which converge at around 20°N (Johns *et al.* 1999). This eddy-constellation could
theoretically transport larvae travelling either north or south along the Saudi Arabian
cost towards Sudan and Eritrea. While this oceanographic feature seems like a
conceivable explanation of the observed patterns, it does not explain why other, similar
eddy formations towards the north do not seem to create similar genetic patterns. In fact,
eddy activity is more pronounced in the northern Red Sea (Zhan et al. unpublished data).
Furthermore, according to recent oceanographic studies in the area, the described gyres
are not as stable as previously assumed (Bower et al. in press) and oceanographic
patterns in the Red Sea in general are not constant over time (P. Zhan, unpublished data).
Before physical models of the area are refined to reflect temporal patterns as well as
circulation in the shallow southern shelf, we feel that inclusion of oceanography at this
point might obscure results.

Nevertheless, the predictive power of our IBE model - despite incorporating only a
single environmental predictor - exceeds that of most previous comparable studies. Using
regression analysis, Alberto et al. (2010) were able to improve their IBD model from
explaining 33% of genetic variability in the giant kelp, Macrocystis pyrifera, to 50% with
the incorporation of habitat discontinuity. A year later Alberto et al. (2011) further
improved their model by incorporating minimum oceanographic distance to explain a
total of 67% of variation. Selkoe et al. (2010) included three predictors of genetic
structure (kelp deb size, sea surface temperature and migration probability) into their
analysis of three species (kelp bass Paralabrax clathratus, Kellet’s whelk Kelletia kelletii
and California spiny lobster Panulirus interruptus) in the Southern California Bight. The
investigators achieved model fits of up to 80% (adjusted $R^2 = 0.8$ for kelp bass),
depending on the combination of predictors involved. Here we explain 92% of genetic
variability using MMRR and 71% using a more conservative approach (partial Mantel tests).

Despite the measurable genetic discontinuity at 19°N, gene flow in *A. bicinctus* along much of the Saudi Arabian coast of the Red Sea seems to follow a stepping-stone mode of dispersal. The stepping-stone pattern becomes most obvious from our network analysis approach, which generated a minimum spanning tree that notably conforms to the geographic layout of the sampling locations without any prior spatial parameter input (Fig. 3). Overall, mean connectivity thus appears to be highest among neighboring locations. STRUCTURE analysis also matches these patterns, showing a gradual shift in cluster assignments along the axis of the Red Sea (Fig. 3).

Network analysis further indicates a center of highest node connectedness in the central Red Sea (Fig. 3). Conceptually decreasing connectedness towards the periphery of the basin is a sensible pattern, yet it might also be an artifact of our sampling design with larger numbers of adjacently located sampling stations in the central region (Fig. 1). Overall, to capture the actual processes of gene flow in the Red Sea it is clearly necessary to include samples from the western coast of the basin. However, permit and visa issues have so far complicated collections in these regions.

**Conclusions:**

Accumulating evidence in marine population connectivity research is pointing towards the critical influence of environmental heterogeneity on patterns of realized larval dispersal (e.g. Marshall *et al.* 2010; Selkoe *et al.* 2010; Shima & Swearer 2010; Burgess
et al. 2012). Here we show concurring data that habitat characteristics can affect patterns of connectivity in marine organisms with pelagic larval dispersal. Biophysical dispersal models should explicitly integrate environmental patchiness for maximum accuracy and future population genetic studies should incorporate seascape features into the analysis.

With plans underway to establish spatial management along the Saudi Arabian Red Sea coast, it is vital to describe patterns of connectivity in the region. Here we show that gene flow in a coral reef fish follows a stepping-stone model along much of the Saudi Arabian coast. Due to a significant genetic break at around 19°N, we recommend separate management of the Farasan region, which may have to rely on internal replenishment due to limited gene flow into this region.
References:


Appendix A: Red Sea bathymetry (a) and mean chl-a concentrations (b). Figures adopted from Raitos et al. (2013), with permission from D. Raitos.
### Appendix B: Alternative measures of pairwise genetic distances among populations – standardized $F_{ST}$ ($F'_{ST}$) below diagonal, Nei’s distance above diagonal.

<table>
<thead>
<tr>
<th></th>
<th>Bur</th>
<th>AN</th>
<th>Nuw</th>
<th>Mas</th>
<th>Amat</th>
<th>Yan</th>
<th>SB</th>
<th>Hai</th>
<th>PR</th>
<th>AM</th>
<th>ER</th>
<th>AT</th>
<th>TR</th>
<th>UH</th>
<th>Can</th>
<th>AQ</th>
<th>Mag</th>
<th>Sum</th>
<th>Far</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bur</td>
<td>0.001</td>
<td>0.004</td>
<td>-0.003</td>
<td>0.005</td>
<td>0.019</td>
<td>0.01</td>
<td>0.011</td>
<td>0.01</td>
<td>0.012</td>
<td>0.011</td>
<td>0.024</td>
<td>0.027</td>
<td>0.015</td>
<td>0.021</td>
<td>0.04</td>
<td>0.062</td>
<td>0.089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN</td>
<td>0.005</td>
<td>-0.004</td>
<td>0.007</td>
<td>0.006</td>
<td>0.023</td>
<td>0.013</td>
<td>0.015</td>
<td>0.009</td>
<td>0.011</td>
<td>0.011</td>
<td>0.013</td>
<td>0.03</td>
<td>0.026</td>
<td>0.018</td>
<td>0.019</td>
<td>0.034</td>
<td>0.056</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>Nuw</td>
<td>-0.004</td>
<td>0.007</td>
<td>0.001</td>
<td>0.015</td>
<td>0.001</td>
<td>0.006</td>
<td>0.006</td>
<td>0.004</td>
<td>0.001</td>
<td>0.006</td>
<td>0.015</td>
<td>0.015</td>
<td>0.006</td>
<td>0.009</td>
<td>0.027</td>
<td>0.047</td>
<td>0.072</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amat</td>
<td>0.006</td>
<td>0.007</td>
<td>0.001</td>
<td>-0.002</td>
<td>0.001</td>
<td>-0.001</td>
<td>0.003</td>
<td>0</td>
<td>0.009</td>
<td>0.008</td>
<td>0.002</td>
<td>0.008</td>
<td>0.012</td>
<td>0.032</td>
<td>0.054</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yan</td>
<td>0.022</td>
<td>0.018</td>
<td>0.026</td>
<td>0.015</td>
<td>0.014</td>
<td>0.018</td>
<td>0.008</td>
<td>0.005</td>
<td>0.01</td>
<td>0.005</td>
<td>0.006</td>
<td>0.017</td>
<td>0.012</td>
<td>0.012</td>
<td>0.016</td>
<td>0.02</td>
<td>0.028</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>SB</td>
<td>0.013</td>
<td>0.007</td>
<td>0.017</td>
<td>0</td>
<td>0</td>
<td>0.005</td>
<td>0.001</td>
<td>0</td>
<td>0.002</td>
<td>0.003</td>
<td>0.009</td>
<td>0.007</td>
<td>0.003</td>
<td>0.009</td>
<td>0.009</td>
<td>0.031</td>
<td>0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hai</td>
<td>0.015</td>
<td>0.009</td>
<td>0.02</td>
<td>0.006</td>
<td>-0.003</td>
<td>0.007</td>
<td>0.006</td>
<td>-0.003</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.002</td>
<td>0.006</td>
<td>0.003</td>
<td>0</td>
<td>0.003</td>
<td>0.008</td>
<td>0.021</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>PR</td>
<td>0.013</td>
<td>0.002</td>
<td>0.012</td>
<td>0.007</td>
<td>0</td>
<td>0.002</td>
<td>0</td>
<td>-0.004</td>
<td>-0.005</td>
<td>-0.005</td>
<td>-0.002</td>
<td>0.011</td>
<td>0.006</td>
<td>-0.001</td>
<td>0.005</td>
<td>0.007</td>
<td>0.023</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>AM</td>
<td>0.012</td>
<td>0.003</td>
<td>0.014</td>
<td>0.004</td>
<td>-0.003</td>
<td>0.009</td>
<td>-0.001</td>
<td>-0.003</td>
<td>-0.008</td>
<td>-0.003</td>
<td>-0.004</td>
<td>0.002</td>
<td>0.002</td>
<td>-0.004</td>
<td>0.007</td>
<td>0.005</td>
<td>0.027</td>
<td>0.037</td>
<td></td>
</tr>
<tr>
<td>ER</td>
<td>0.015</td>
<td>0.008</td>
<td>0.014</td>
<td>0</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>-0.002</td>
<td>-0.007</td>
<td>-0.006</td>
<td>-0.006</td>
<td>0.005</td>
<td>0.001</td>
<td>-0.002</td>
<td>0.006</td>
<td>0.011</td>
<td>0.024</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td>AT</td>
<td>0.014</td>
<td>0.005</td>
<td>0.017</td>
<td>0.007</td>
<td>-0.001</td>
<td>0.006</td>
<td>0.004</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.006</td>
<td>-0.008</td>
<td>-0.001</td>
<td>0</td>
<td>-0.003</td>
<td>0.001</td>
<td>0.005</td>
<td>0.018</td>
<td>0.038</td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>0.029</td>
<td>0.023</td>
<td>0.037</td>
<td>0.016</td>
<td>0.009</td>
<td>0.016</td>
<td>0.009</td>
<td>0.005</td>
<td>0.012</td>
<td>0</td>
<td>0.004</td>
<td>-0.002</td>
<td>0.001</td>
<td>-0.002</td>
<td>0.006</td>
<td>0.003</td>
<td>0.015</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>UH</td>
<td>0.035</td>
<td>0.022</td>
<td>0.033</td>
<td>0.018</td>
<td>0.009</td>
<td>0.012</td>
<td>0.008</td>
<td>0.002</td>
<td>0.006</td>
<td>0.001</td>
<td>0.001</td>
<td>0</td>
<td>-0.001</td>
<td>-0.005</td>
<td>0.002</td>
<td>-0.001</td>
<td>0.015</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td>Can</td>
<td>0.019</td>
<td>0.006</td>
<td>0.022</td>
<td>0.005</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
<td>-0.002</td>
<td>-0.003</td>
<td>-0.007</td>
<td>-0.004</td>
<td>-0.004</td>
<td>-0.006</td>
<td>-0.008</td>
<td>-0.001</td>
<td>-0.002</td>
<td>0.014</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td>AQ</td>
<td>0.027</td>
<td>0.017</td>
<td>0.025</td>
<td>0.011</td>
<td>0.009</td>
<td>0.018</td>
<td>0.011</td>
<td>0.003</td>
<td>0.006</td>
<td>0.008</td>
<td>0.007</td>
<td>0.001</td>
<td>0.005</td>
<td>0.001</td>
<td>-0.003</td>
<td>0.005</td>
<td>0.015</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>Mag</td>
<td>0.053</td>
<td>0.034</td>
<td>0.046</td>
<td>0.035</td>
<td>0.015</td>
<td>0.021</td>
<td>0.011</td>
<td>0.011</td>
<td>0.009</td>
<td>0.006</td>
<td>0.014</td>
<td>0.007</td>
<td>0.001</td>
<td>-0.003</td>
<td>-0.004</td>
<td>0.006</td>
<td>0.005</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>0.083</td>
<td>0.065</td>
<td>0.076</td>
<td>0.063</td>
<td>0.043</td>
<td>0.035</td>
<td>0.042</td>
<td>0.028</td>
<td>0.033</td>
<td>0.036</td>
<td>0.034</td>
<td>0.027</td>
<td>0.019</td>
<td>0.019</td>
<td>0.019</td>
<td>0.021</td>
<td>0.005</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>Far</td>
<td>0.119</td>
<td>0.095</td>
<td>0.12</td>
<td>0.099</td>
<td>0.075</td>
<td>0.09</td>
<td>0.063</td>
<td>0.059</td>
<td>0.061</td>
<td>0.053</td>
<td>0.063</td>
<td>0.055</td>
<td>0.039</td>
<td>0.044</td>
<td>0.039</td>
<td>0.052</td>
<td>0.023</td>
<td>0.018</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Log-likelihood of the number of population clusters explaining the genetic data showing the average values of $\ln P(K)$ resulting from a minimum of 10 runs ($\pm 1$SE).
Not finding Nemo: empirical evidence of an open reef fish population

Abstract: The degree to which marine populations are open (connected to other subpopulations), or closed (mostly self-seeding) is of fundamental importance for the demographic and evolutionary dynamics of species with pelagic larval stages (Mora & Sale 2002; Kritzer & Sale 2004; Cowen & Sponaugle 2009). Recently emerging empirical evidence from studies of demographic connectivity has created a paradigm shift towards an assumption that populations are more closed than previously believed (Jones et al. 1999; Swearer et al. 1999; Paris & Cowen 2004; Gerlach et al. 2007; Buston et al. 2011; Almany et al. 2013). Measurements of larval dispersal in various species and locations have consistently found high levels of self-recruitment (SR) in coral reef fish, usually within the order of 20-70% (Jones et al. 2009; Berumen et al. 2012). Here we show the first empirical evidence that some coral reef fish populations persist by receiving recruits exclusively from external sources, lacking any demographically relevant SR. We conducted genetic parentage analysis on the anemonefish, *Amphiprion bicintus* (Rüppel 1830), at a coral reef in the Red Sea and found 1% and 0% self-recruitment in two consecutive years. Juvenile collections at surrounding reefs resulted in one extra parent-offspring match in 2013. We additionally ran a biophysical dispersal model of the study system, parameterized with real-time climatological forcing. Model predictions concurred entirely with the empirical data. Our results demonstrate that populations of coral reef fishes can be virtually open on scales of several kilometers.
These findings introduce an important exception to the emerging paradigm of larval connectivity, and present an urgent caution when assuming self-replenishment in marine reserve design.

**Introduction:**

Traditionally, marine organisms with planktonic larval life history phases were considered to be largely open with rates of recruitment decoupled from local production (Roberts 1997). Over the past 15 years however, a new picture has been emerging, assuming populations to exist in a variable continuum between open and closed states (Jones et al. 2009; Saenz-Agudelo et al. 2012; Pinsky et al. 2012). Especially in the study of coral reef fishes, the contemporary paradigm emphasizes the importance of native production to local replenishment. Recent technologies (e.g. larval tagging, parentage analysis) have facilitated empirical estimates of demographic patterns of larval dispersal. The vast majority of studies to date that have employed such approaches to empirically assess dispersal patterns in coral reef fishes have yielded surprisingly high levels of self-recruitment (Table 1). Moreover, recruitment of propagules was shown to decrease rapidly with increasing distances from the natal site (Buston et al. 2011; Almany et al. 2013; D’Aloia et al. 2013). These findings have vast consequences for the spatial management of marine populations in that locally produced recruits could potentially offset a lack of externally derived settlers. In other words, if SR was indicative of the potential for self-replenishment, then recent findings suggest that the appropriate spatial scale of individual MPAs would be much smaller than historically assumed, because
<table>
<thead>
<tr>
<th>Fish family</th>
<th>PLD</th>
<th>Location</th>
<th>Scale [km²]</th>
<th>Distance [km]</th>
<th>Method</th>
<th>Adults sampled</th>
<th>Proportion [%]</th>
<th>Recruits sampled</th>
<th>Recruits assigned</th>
<th>%SR</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pomacentridae</td>
<td>24</td>
<td>GBR</td>
<td>4.0</td>
<td>5.0</td>
<td>TCT</td>
<td>n/a</td>
<td>0.5 - 2*</td>
<td>5000</td>
<td>15</td>
<td>15-60</td>
<td>(Jones et al. 1999)</td>
</tr>
<tr>
<td>Amphiprioninae</td>
<td>11</td>
<td>PNG</td>
<td>0.5</td>
<td>1.0</td>
<td>TCT/Par.</td>
<td>55</td>
<td>100</td>
<td>136</td>
<td>33</td>
<td>16-32</td>
<td>(Jones et al. 2005)</td>
</tr>
<tr>
<td>Amphiprioninae</td>
<td>11</td>
<td>PNG</td>
<td>0.8</td>
<td>6.0</td>
<td>BT</td>
<td>176</td>
<td>100</td>
<td>15</td>
<td>9</td>
<td>60</td>
<td>(Almany et al. 2007)</td>
</tr>
<tr>
<td>Chaetodontidae</td>
<td>38</td>
<td>PNG</td>
<td>0.8</td>
<td>6.0</td>
<td>BT</td>
<td>123</td>
<td>17.3</td>
<td>77</td>
<td>8</td>
<td>60</td>
<td>(Almany et al. 2007)</td>
</tr>
<tr>
<td>Amphiprioninae</td>
<td>11</td>
<td>PNG</td>
<td>0.8</td>
<td>6.0</td>
<td>Par.</td>
<td>253</td>
<td>100</td>
<td>254</td>
<td>107</td>
<td>42</td>
<td>(Planes et al. 2009)</td>
</tr>
<tr>
<td>Amphiprioninae</td>
<td>11</td>
<td>PNG</td>
<td>17</td>
<td>1.2</td>
<td>Par.</td>
<td>451</td>
<td>100</td>
<td>491</td>
<td>89</td>
<td>18</td>
<td>(Saenz-Agudelo et al. 2011)</td>
</tr>
<tr>
<td>Amphiprioninae</td>
<td>11</td>
<td>PNG</td>
<td>0.8</td>
<td>6.0</td>
<td>Par.</td>
<td>267</td>
<td>100</td>
<td>161</td>
<td>103</td>
<td>64</td>
<td>(Berumen et al. 2012)</td>
</tr>
<tr>
<td>Chaetodontidae</td>
<td>38</td>
<td>PNG</td>
<td>0.8</td>
<td>6.0</td>
<td>Par.</td>
<td>175</td>
<td>22</td>
<td>103</td>
<td>9</td>
<td>40</td>
<td>(Berumen et al. 2012)</td>
</tr>
<tr>
<td>Serranidae</td>
<td>26</td>
<td>GBR</td>
<td>2.0</td>
<td>0.0</td>
<td>Par.</td>
<td>233</td>
<td>100</td>
<td>140</td>
<td>17</td>
<td>7</td>
<td>(Harrison et al. 2012)</td>
</tr>
<tr>
<td>Lutjanidae</td>
<td>25</td>
<td>GBR</td>
<td>2.0</td>
<td>0.0</td>
<td>Par.</td>
<td>577</td>
<td>100</td>
<td>171</td>
<td>23</td>
<td>19</td>
<td>(Harrison et al. 2012)</td>
</tr>
<tr>
<td>Gobiidae</td>
<td>21</td>
<td>MBR</td>
<td>0.25</td>
<td>1.0</td>
<td>Par.</td>
<td>212</td>
<td>100</td>
<td>194</td>
<td>9</td>
<td>4.6</td>
<td>(D’Aloia et al. 2013)</td>
</tr>
<tr>
<td>Serranidae</td>
<td>25</td>
<td>PNG</td>
<td>5.0</td>
<td>0.0</td>
<td>Par.</td>
<td>235</td>
<td>43</td>
<td>204</td>
<td>27</td>
<td>17-25</td>
<td>(Almany et al. 2013)</td>
</tr>
<tr>
<td>Amphiprioninae</td>
<td>11</td>
<td>Red Sea</td>
<td>0.7</td>
<td>4.0</td>
<td>Par.</td>
<td>150</td>
<td>99</td>
<td>171</td>
<td>1</td>
<td>0.5</td>
<td>this study</td>
</tr>
</tbody>
</table>

**Table 1:** Summary statistics of studies estimating self-recruitment using comparable approaches to this study displaying: pelagic larval durations (mean estimates); study locations: GBR = Great Barrier Reef, PNG = Papua New Guinea, MBR = Mesoamerican Barrier Reef; study scale: area of self-recruitment estimates; minimum distance of closest habitat of study species beyond the study area; assignment method: TCT = Tetracycline tagging, Par. = Parentage analysis, BT = Barium tagging; total number of females sampled (if not provided by the authors, total adults was divided by two); approximate proportion of adults sampled inside area of self-recruitment estimates (as indicated by the authors); total number of recruits collected for self-recruitment estimates over entire study period; total number of recruits identified as natal recruits; estimated percent self-recruitment inside study area. *Proportion of eggs marked
populations regularly exhibit high levels of SR. Here we challenge this paradigm by showing that demographic dynamics in coral reef metapopulations may be characterized by the presence of perfect sink populations.

**Methods:**

We used genetic parentage analysis to measure levels of SR and dispersal to surrounding reefs in an anemonefish, *Amphiprion bicinctus*, at a reef, Quita al Girsh (QG; 22.431°N, 38.995°E), on the Saudi Arabian coast of the Red Sea (Fig.1). QG was chosen as the focal reef after an extensive survey of the area because of its isolated location (closest reef separated by 4 km of water in excess of 400 m depth) and its appropriate spatial and ecological scale (c. 1 x 0.4 km, c. 160 anemones).

**Sampling:**

During a thorough and systematic search, involving up to three dive teams, over a one-month period, all anemones (*n* = 152) at QG were geolocated using time-synchronized, surface-towed GPS units and marked with individual tags and visual markers to facilitate subsequent recovery (Appendix A). All encountered adult anemonefish were collected using hand nets, measured to the nearest millimeter, tail clipped and released back into their host anemone. All samples were preserved in 96% ethanol and transported to the laboratory for subsequent genetic analyses. In December and January 2012, we collected tissue samples of 311 potential parents from 152 anemones at QG, accounting for approximately 99% of all adult individuals at that site.
Figure 1: Location and abundance of sampled juveniles. The purple circle indicates Quite al Girsh (QG), the location of adult sampling (311 potential parents) with the according total number of juveniles collected in 2012 and 2013 (353 juveniles). Red circles are scaled to the number of juveniles collected at surrounding reefs in 2012 and 2013 combined. Arrows indicate the locations of juvenile parentage assignments.

Prior to this study no published data on spawning seasonality in A. bicinctus was available and even veteran Red Sea explorers did not have reliable information (e.g. F. Krupp pers. comm.). We first observed a spawning event in January 2010 during collections for Chapter 4, which was marked by the presence of recent recruits in the majority of sampled anemones. The first collections of juveniles for parentage assignment analysis were accordingly carried out in January 2012 approximately two
weeks after the full moon, which is the spawning time of most related species (e.g. Ross 1978). Subsequent surveys in April, July, September, and November 2012 revealed occasional eggs and/or juveniles after the full moon, but no mass spawning events. The next observed spawning cycle happened in January 2013, the time of the second round of collections.

During collections in January 2012 and January and February 2013, up to three dive teams searched every marked anemone at QG for recent recruits (< 2 cm), which were collected using hand nets, anaesthetized using clove oil and retained in ethanol. We collected 71 and 101 recent recruits in 2012 and 2013 from QG respectively (Fig.1). Collections at surrounding reefs in both years were carried out in a similar manner. Recent recruits were sampled opportunistically during SCUBA dives aimed at covering large areas of each respective reef. Anemones containing juveniles were geolocated as per QG, but no physical markers were placed. We collected 16 and 160 juveniles in 2012 and 2013 respectively at reefs surrounding QG (Fig.1).

**Genetic analysis:**

In total, we screened a total of 659 individuals at 30 microsatellite loci (Nanninga et al. 2012, Chapter 3). The data was tested for the presence of null-alleles using Microchecker (v2.2.3; van Oosterhout et al. 2004). Deviations from Hardy-Weinberg-Equilibrium (HWE) and Linkage Disequilibrium (LD) between loci were tested in GENEPOP (web-service; Raymond & Rousset 1995), adjusting for multiple tests using False Discovery Rates (FDR, Benjamini & Hochberg 1995). Unbiased heterozygosities (uH), and mean
within-group pairwise relatedness ($r$) were calculated in Genalex (v6.5; Peakall & Smouse 2012).

Using FAMOZ we identified putative parent-offspring pairs in our data set. FAMOZ was shown to produce reliable kinship estimates if sufficient polymorphic loci are employed (Harrison et al. 2013). The software employs a categorical allocation algorithm based on a maximum-likelihood approach to compute LOD scores for assigning individuals to candidate parents based on the observed allelic frequencies at each locus (Gerber et al. 2003; Harrison et al. 2013). We simulated 10,000 parent-offspring pairs based on observed allele frequencies and another 10,000 pairs from the putative parental genotypes. Minimum LOD score thresholds to accept single parent- or parent-pair-offspring assignments were defined by identifying the intercept of the frequency distributions of the two simulations. Introduction of error (0.1% and 1%) did not alter the results.

*Dispersal model*

Moreover, we ran a biophysical dispersal model simulating larval trajectories of *A. bicinctus* within the study system. The Connectivity Modeling System (CMS) is a probabilistic model of larval dispersal integrating physical forcing with a stochastic Lagrangian framework that incorporates biological attributes (Paris et al. 2013). We ran the model with averaged climatological forcing for the month of collections (January) and biological attributes (i.e. vertical migration probabilities) estimated from the
literature of closely related species (Irisson et al. 2010; Huebert et al. 2011). Virtual particles were released from QG and allowed to disperse for a maximum of 12 days.

The physical features of the dispersal model were based upon an MIT general circulation model (MITgcm). The MITgcm provides horizontally uniform vertical resolution, thus decreasing computational error due to steep bottom topography. We employed a small-domain, high-resolution model, ranging from 22.1 to 22.8°N and 38.5 to 39.2°E and with a grid size of 0.005°. The model was nested by velocity, temperature and salinity on north, west and south boundaries from a coarser model of the entire Red Sea simulated over 50 years. A climatological dataset restoring 3-day averaged output provided the boundary conditions for the small-domain model.

The physical current data provided the background hydrodynamics of the CMS, which calculated particle trajectories based on Lagrangian descriptions of oceanic phenomena. Release and settlement sites were based on reef habitat polygons with a grid size of 500 m². The CMS includes a mortality algorithm based upon maximum pelagic larval duration (PLDmax = 12 days). Larvae were considered to have settled if they were retained by a polygon after a minimum of eight days. We released 8,000 particles, which was shown to be sufficient to reduce the fraction of unexplained variance (Simons et al. 2013), and ran the model with three different vertical distribution files with different maximum and mean depths. The results did not differ among model runs.
Estimation of mean dispersal distance

Mean dispersal distance was estimated from the slope of an isolation-by-distance function based on population genetic data along 1500 km of the Saudi Arabian Red Sea coast (see Chapter 4). The slope is related to the inverse of $4D_e\sigma^2$, where $D_e$ refers to the effective density of individuals and $\sigma$ refers to dispersal spread, the standard deviation of parental versus offspring position (Rousset 1997). Dispersal spread can be calculated by substituting values of slope and appropriate effective densities ($D_e$). Estimates of $\sigma$ can then be related to mean dispersal distance $d$, with $\sigma = 2d$ (Pinsky et al. 2010). Average effective densities were estimated to range between 50 and 400 individuals per km$^2$, based on abundances at GQ.

Results:

All loci satisfied Hardy-Weinberg-Equilibrium assumptions after corrections for false discovery rates. Missing values accounted for 0.3% of the total data set. The mean number of alleles over all loci was 15.9 (7-31) and observed heterozygosity was $0.76 \pm 0.109$ SD (Appendix B).

Mean unbiased heterozygosity did not differ significantly among the adult and recruit cohorts (ANOVA, $p = 0.78$), neither did within-cohort pairwise relatedness (ANOVA, $p = 0.32$) (Fig. 2).
Figure 2: Genetic group statistics for all adults from QG, all recruits from 2012 and 2013. Mean unbiased heterozygosity (a) and mean within-group pairwise relatedness ($r$) ± bootstrapping error (b).

Out of a total of 348 juveniles sampled (87 in 2012; 261 in 2013), two individuals were assigned to parents at QG (Fig. 1). One individual recruited back to QG in 2012, the other one traveled approximately 10 km north to settle at an adjacent reef in 2013 (Fig. 1). At the reef scale of QG (~0.7 km$^2$), this equates to approximately 1% self-recruitment in 2012 and 0% in 2013. One recruit was assigned to a parent pair within one anemone; the other was a single parent assignment.

Out of 8,000 virtual larvae released by the CMS, only three settled inside the study domain (Fig. 3). All remaining particles died or left the study area within ten days after release. The retained individuals settled at a reef north of QG, in close vicinity of the location of our only empirical assignment from 2013 (Fig. 3).
Figure 3: Connectivity Modeling System simulations of virtual larvae released from QG (green star). Red particles indicate trajectories of larvae that have left the model domain after 10-12 days. Blue particles indicate trajectories of larvae that settled (n = 3, location: blue arrow). Black arrow: location of parentage assignment.

Estimates of dispersal distance ranged from 5.6 to 15.8 km (mean = 9.4 km).

Spawning in A. bicinctus in the central Red Sea follows a strongly seasonal pattern with the bulk of reproduction occurring in the winter months (January and February) and only sporadic recruitment during the rest of the year.
Discussion:

Our results indicate that, despite recently accumulating evidence for high levels of self-recruitment, there is no general pattern of high local retention in coral reef fishes and populations can be effectively open at scales of tens of kilometers.

Conceptually, open marine populations are clearly not a novel notion (Roughgarden et al. 1988), yet this is the first empirical demonstration of this concept in a coral reef fish. Furthermore, we present the first integration of empirical connectivity estimates with a biophysical dispersal model of the same system. Such an approach has been advocated repeatedly in the recent marine ecological literature (e.g. Werner et al. 2007; Botsford et al. 2009; Leis et al. 2011a), yet as so far not been implemented in a published study. This study produced a substantial data set, the likes of which only exist in a hand-full of studies worldwide (Table 1). Our results are particularly remarkable when compared to the findings of these existing empirical dispersal studies. While our spatial scale, sampling regime and study species are comparable to earlier investigations, the percentage of assignments was much lower than expected (Table 1). The only other comparable study that has found similarly low values of SR was Saenz-Agudelo et al. (2011, 2012), who estimated 0-2% SR in a small sub-population of the panda clownfish, *Amphiprion polymnus*. The difference in population size, however, exceeds an order of magnitude (15 versus 152 anemones) and is thus not on par with our findings.

As both our empirical and modeling approaches indicate, all larvae were transported out of the study domain (parentage: ~30 km$^2$, CMS: ~67 km$^2$) well before the end of the pelagic larval phase. The congruence between our empirical and simulated data is
remarkable and underlines the validity of our results. Further support for an open system stems from our population genetic analyses. Differences in heterozygosity among cohorts might imply sweepstakes events, whereby only a few adults produce the bulk of the recruits of a generation (Christie et al. 2010). Alternatively, it could indicate high levels of self-recruitment, where the adults at a particular location predominantly produce the next local generation. Here we found no difference in genetic diversity among cohorts (Fig. 2a), indicating that recruits at our sampling locations originate from a widespread group of adults. Similarly, we detected low pairwise relatedness within cohorts, which did not differ among cohorts (Fig. 2b). This pattern further indicates that recruits arrived from a range of different reefs rather than from a restricted number of adults.

Overall, our results stand in a stark contrast to previous hypotheses of widespread biophysical retention mechanisms (Swearer et al. 1999, 2002; Sponaugle et al. 2002; Paris & Cowen 2004). Moreover, closely related species were shown to exhibit behavioural patterns that may promote self-recruitment (Gerlach et al. 2007; Leis et al. 2011b) and retention in the natal habitat may entail selective advantages due to decreased costs of dispersal (Nosil et al. 2005; Marshall et al. 2010; Bonte et al. 2012; also see Chapter 6). Yet this does not seem to hold true for A. bicinctus at our focal resolution.

Characteristic current velocities in our study domain were approximately 5 cms$^{-1}$, less than most mean transport speeds reported for different regions around the globe and approximately four times less than reported swimming speeds of closely related species (Fisher et al. 2005). A. bicinctus should hence be able to influence their trajectories at least to some degree. Yet again, flow was remarkably laminar with little eddy formation to entrap larvae around reefs.
Mean dispersal distance was estimated to range between 5.6 and 15.8 km with a mean of 9.4 km. These values concur with previous estimates for related species (Pinsky et al. 2010) and do not explain the observed patterns in relation to previous self-recruitment studies.

Recent research has highlighted the importance of the geographic setting in determining levels of self-recruitment (Saenz-Agudelo et al. 2011; Pinsky et al. 2012; D’Aloia et al. 2013). Due to dilution effects from nearby patches, measures of self-recruitment may become attenuated in continuous habitats (Pinsky et al. 2012). Despite the continuous nature of Red Sea fringing reefs however, our results undermatch any comparable previous studies with even closer adjacent habitats (Table 1) and our model simulations confirm zero self-recruitment at QG.

Hypothetically, our findings might instead indicate an active choice to passively drift. If local retention is an active behaviour (e.g. Gerlach et al. 2007), then so might be dispersal. The propensity of individuals towards dispersal in different terrestrial species was shown to be genetically inherited (Ahlroth et al. 2010) or maternally induced (Tschirren et al. 2007). The tendency of marine larvae to return to the natal patch or disperse away might accordingly evolve as a function of increases in habitat patchiness (Baskett et al. 2007) or local conditions (Bowler & Benton 2011). Future research should consider the existence of such ‘dispersotypes’ (Ronce & Clobert 2012) when interpreting patterns of larval dispersal (see Chapter 6).

Our findings may have important ramifications for the design of spatial management strategies. We showed that at least in some locations and for some species, populations of
coral reef fishes might be virtually open, i.e., devoid of SR. With no local recruitment, many populations might not be able to sustain themselves in the absence of substantial levels of immigration.
130

References:


Appendix A: Locations of anemones around Quita al Girsh.
Appendix B: Numbers of alleles (bars), observed (grey lines) and expected (black lines) heterozygosities across all loci. GenBank accession numbers available from Nanninga et al. (2012).
There’s no place like home – origin effects in marine population connectivity

Abstract: Marine metapopulation dynamics may be regulated by local replenishment of sedentary adult populations via pelagic larvae produced by local and/or external sources. Most populations are assumed to exist in a continuum between self-seeding and inter-patch connectivity, the ratio of which might be highly variable in space and time. Emerging evidence from empirical measurements of larval dispersal indicates high levels of self-recruitment and dispersal kernels describing rapidly decreasing probability functions. The recent focus on establishing patterns of connectivity has focused primarily on settlement connectivity, raising important discussions about the degree to which observed settlement translates into reproductive individuals making contributions to future generations. For such realized connectivity to take place, individuals need not only to successfully settle into a location, but also survive and reproduce. Early life history stages in most marine organisms are characterized by high rates of mortality; patterns of larval dispersal may therefore not reflect realized connectivity due to selective post-settlement processes. Mortality schedules in many teleost fish and invertebrates appear to be closely linked to particular phenotypic traits, such as growth, size, and/or overall condition. Recruit phenotypes in turn might be strongly influenced by carry-over effects from embryonic and larval stages through variable parental effects and different dispersal histories. The cost of dispersal theoretically increases with distance between natal and settlement sites, yet the cost-benefit ratio will vary intra-s Specifically depending
on a combination of parental and environmental effects, individual phenotype, and habitat fragmentation. To conceptualize these recent and resurgent theories into a comprehensive hypothesis, we put forward the concept of origin effects, which we briefly define as the fragmentation-dependent influence of natal origin and dispersal history on individual phenotypes and the effect thereof on post-settlement selection. We discuss the potential influence of origin effects on patterns of larval connectivity and set it in context with recent findings.

**Introduction:**

Over the past two decades, increasing attention has been paid to the role that larval connectivity plays in population dynamics in marine organisms (Sale et al. 2006; Jones et al. 2009; Jones et al. 2009; Hixon 2011). Many marine systems are characterized by patchily distributed, relatively sedentary adult populations that are interconnected to varying degrees via a pelagic larval stage of species-specific duration (days to months) (Kritzer & Sale 2004). In a metapopulation context, connectivity is hence defined as the exchange of individuals between spatially discrete subpopulations via larval dispersal (Sale et al. 2006; Pineda et al. 2007). The extent to which local populations exchange, receive, and/or self-seed recruits can have important ramifications for population dynamics from demographic (population growth, persistence, resilience) to evolutionary scales (gene flow, adaptation) and should influence the way we manage and conserve coastal populations (Cowen & Sponaugle 2009; Gaines et al. 2010).
It is currently assumed that most marine metapopulations exist somewhere in a continuum between closed (only self-seeding) and open (only externally seeded) systems, the measure of which may be spatially and/or temporally variable (Jones et al. 2009, Berumen et al. 2012). Despite indications for genetic panmixia over large geographic scales and long distance dispersal events in excess of 100s km (e.g. Shulman & Bermingham 1995; Craig et al. 2007; Horne et al. 2008), there is accumulating evidence for positively skewed dispersal kernels (Buston et al. 2011; Horne et al. 2011; Salinas-de-León et al. 2012) and surprisingly high fractions of self-recruitment in several species of invertebrates and teleost fishes (e.g. Swearer et al. 2002; Taylor & Hellberg 2003; Cowen et al. 2006; Planes et al. 2009). These patterns seem to be largely independent of differences in life histories (Almany et al. 2007; Shanks 2009; Carson et al. 2011) and can be remarkably constant over time (Berumen et al. 2012; Priest et al. 2012; Saenz-Agudelo et al. 2012; but see Chittaro & Hogan 2012). A recent hypothesis emphasizes the importance of the geographic setting, or habitat patchiness in determining the rates of larval exchange and self-seeding (Jones et al. 2009; Saenz-Agudelo et al. 2011; Andutta et al. 2012; Pinsky et al. 2012). This theory is substantiated by remarkably similar empirical estimates of connectivity patterns of different species (and even taxa) in shared habitats (Almany et al. 2007; López-Duarte et al. 2012) versus apparent variations in patterns within one species between two different habitats (Jones et al. 2005; Saenz-Agudelo et al. 2011). Pinsky et al. (2012) recently introduced a model that predicts levels of self-recruitment to be positively correlated with habitat patchiness, potentially explaining the existence of relatively closed populations despite high dispersal potential. Isolated patches would thus theoretically exhibit higher rates of self-recruitment than
patches in denser environments due to higher dilution effects in the latter. If this hypothesis is correct, the question remains whether there is reduced dilution in isolated patches because of a shortage of external recruits that arrive at the patch or because dispersal costs for long-distance external recruits result in post-settlement selection against immigrants (Marshall et al. 2010; Bonte et al. 2012; Burgess et al. 2012). An alternate hypothesis is offered by, Andutta et al. (2012), predicting high rates of retention in densely aggregated reef areas due to reduced water flow and higher residence times of passive particles. We argue that both theories are not mutually exclusive and that rates of local retention could be decoupled from observed self-recruitment, where different dispersal histories create individual larval and recruit phenotypes that vary in their adaptation to selective post-settlement processes.

In marine organisms with complex life cycles, the dispersal phase represents a potential population bottleneck involving exceedingly high rates of mortality during the larval stage due to predation, starvation, and advection to unsuitable habitats (Houde 1989; Leggett & Deblois 1994). Even small fluctuations in larval mortality can potentially cause large variations in recruit cohorts (Caley et al. 1996; Houde 1997), yet predictions about population dynamics based solely on colonizer numbers may be inadequate (Pineda et al. 2010; Burgess & Marshall 2011). The process of metamorphosis and recruitment into the juvenile and adult population represents another critical bottleneck regularly involving the loss of more than half of the cohort immediately after settlement (Caley 1998; Doherty et al. 2004; Almany & Webster 2006). Successful larval exchange between habitats may thus not effectively reflect
realized connectivity, especially if incoming recruits die non-randomly shortly after settlement.

Early life-history mortality is often selective with respect to specific traits (e.g. Perez & Munch 2010) and “high quality” individuals should theoretically contribute proportionately more to the population than those in poor condition (Suthers 1998). Selective post-settlement processes can thus effectively decouple patterns of larval dispersal from true reproductive connectivity, the fraction of recruits that survives to reproduce (Pineda et al. 2007; Marshall & Morgan 2011; Burgess et al. 2012). Recent work highlighted the importance of larval phenotype in regulating marine populations (e.g. Burgess & Marshall 2011), while at the same time it was demonstrated that juvenile survival may in fact have a greater influence on population growth than do patterns of patch connectivity (Figueira 2009; Carson et al. 2011; López-Duarte et al. 2012). It is becoming increasingly apparent that we need to understand not only the influence of larval quantity reaching a population, but also the variations in larval quality (Allen & Marshall 2010; Burgess & Marshall 2011).

Intraspecific recruit quality can vary notably in space and time (Marshall & Keough 2008), thereby possibly affecting patterns of population dynamics and connectivity (Marshall & Morgan 2011). Recent work has shown that life history stages in many marine organisms are strongly linked and that experiences in one stage can carry over to affect performance in subsequent stages (reviewed in Marshall & Morgan 2011). The origin and larval history of recruiting individuals might hence notably affect patterns of post-settlement survival, thus shaping realized connectivity (Allen & Marshall 2010). Most models of marine larval dispersal and population connectivity treat the dispersal
environment as a binary landscape represented by patches of equal quality separated by a homogenous (yet dynamic) matrix, assigning all simulated recruits an equal chance of survival to reproduction (e.g. Cowen et al. 2006; Paris et al. 2007; Siegel et al. 2008). Although these models undoubtedly present powerful tools to simulate and predict patterns of larval dispersal, the assumption of an unstructured landscape will need to be adapted to recent findings demonstrating the impact of dispersal history on post-settlement processes (Shima et al. 2010). Identifying the phenotypes associated with the selective loss of individuals and the influence of individual backgrounds and experiences on specific traits will be vital for a better understanding of population regulation (Hoey & McCormick 2004).

Here we introduce the concept of origin effects, integrating the relationship between linked life history stages, post-settlement selection and spatial habitat arrangement and their influence on patterns of realized connectivity. We first review the different aspects related to origin effects. In the light of recent findings, we then discuss the theoretical implications of origin effects on metapopulation connectivity. This paper is not meant to be an exhaustive review of all the involved subject matters – there are excellent recent accounts (e.g. connectivity: Cowen & Sponaugle 2009; phenotype-environment mismatches: Marshall et al. 2010; linked life histories: Marshall & Morgan 2011; dispersal costs: Bonte et al. 2012) – but rather serve as a synthesis of recently emerging and resurgent concepts in the regulation of marine metapopulation connectivity that have so far largely gone overlooked.
**Glossary of terms:**

**Carry-over effects:** Traits (e.g. growth rates) and conditions (e.g. fat content), which are present in one life stage (embryonic, larval and/or juvenile) and carry over to subsequent stages (Pechinick 2006).

**Latent effects:** Traits and conditions that originate in one life stage, but only manifest themselves during later stages (Pechinick 2006).

**Parental effects:** The non-genetic influence of parent environment, phenotype, and conditioning on the phenotype of offspring (Green 2008; Marshall et al. 2008).

**Matrix effects:** The influence of the dispersal environment (e.g. temperature, salinity, food availability) on larval condition (Shima & Swearer 2009; 2010).

**Legacy effects:** Carry-over effects of matrix effects (Shima & Swearer 2009; 2010)

**Dispersal costs:** The potential energetic costs and resulting adverse effects of the dispersal process on larval and/or recruit condition. Dispersal costs can be separated into *immediate* (i.e. larval mortality) and *deferred* costs, the latter of which occur after settlement. We differentiate between *direct* and *indirect* deferred costs.

**Direct deferred costs:** The energetic costs of dispersal manifest themselves in reduced body condition during or after settlement. Basically, direct deferred dispersal costs are a result of negative matrix and legacy effects.

**Indirect deferred costs:** Settlement into a habitat that does not suit traits adapted to the natal environment might result in *phenotype-environment mismatches* (DeWitt et al. 1998), rendering external recruits less competitive than local recruits (Nosil et al. 2005). Theoretically, dispersal costs should correlate with dispersal time and/or distance (Rousset & Gandon, 2002).
**Dispersotypes:** Intra-specific phenotypic (morphological, physiological, or behavioral) distinction between individuals with different propensities towards dispersal and/or local retention. Dispersotypes may evolve as a function of habitat fragmentation (Ronce & Clobert 2012).

**Origin effects:** The combination of the above concepts into a comprehensive theory. Origin effects integrate the cumulative effects of natal origin, dispersal history, and settlement location via parental, carry-over, latent, and/or legacy effects on individual phenotypes and the influence thereof on post-settlement selection. Larval and recruit phenotype in turn will influence individual propensities towards dispersal versus retention, depending on habitat fragmentation. Theoretically, origin effects should favour individuals returning to their natal habitat and adverse effects of dispersal should scale with patch distance and habitat heterogeneity.

*Linked life histories and selective post-settlement processes*

In many marine organisms with complex life histories, different life stages are remarkably linked (reviewed in Marshall & Morgan 2011). Parental provisioning as well as experiences during the embryonic and/or larval phase have the potential to carry over to subsequent stages to influence survival and/or performance of individuals (Pechenik 2006). These ‘carry-over’ or ‘latent effects’ can thus have important ramifications for connectivity and population dynamics by decoupling larval dispersal from realized connectivity (Pechenik 2006; Gagliano *et al.* 2007; Marshall & Morgan 2011).
Understanding the influence of larval source and dispersal history on the condition of successful recruits is hence essential for our understanding of population regulation and the proper management of marine populations (Hamilton 2008).

The growth-mortality concept

The conceptual framework that is regularly employed to study the relationship between embryonic/larval traits, recruitment success, and post-settlement survival is provided by the ‘growth-mortality hypothesis’ (Anderson 1988). The theory predicts a competitive advantage of faster larval growth through either condensing the larval phase (Houde 1987), or attaining a larger size at age (Miller et al. 1988). Indeed, offspring size was shown to be an effective predictor of performance over a wide range of taxa (Marshall & Keough 2008) and mortality in larval fish is largely driven by size-selective processes (Meekan & Fortier 1996; Wilson & Meekan 2002). These size advantages can prevail across different life stages and individuals that exhibit larger size-at-hatching and/or faster larval growth rates often show higher rates of relative survival upon settlement (Bergenius et al. 2002; Vigliola & Meekan 2002; McCormick & Hoey 2004; Macpherson & Raventos 2005; Hamilton et al. 2008; Dias & Marshall 2010).

Yet growth-related carry-over effects might not always act on selective processes in one direction, but may be strongly dependent on ontogenetic shifts in selective pressures (Gagliano et al. 2007; Meekan et al. 2010). Growth- and size-selective processes in the marine environment were shown to be highly variable even between closely related species (D’Alessandro et al. 2013), ontogenetically within species (Gagliano et al. 2007;
Dias & Marshall 2010; Kesselring et al. 2012), between microhabitats (Smith & Shima 2011), and depending on environmental heterogeneity and competition (Marshall et al. 2006; Johnson 2008; Johnson & Hixon 2010). Care should hence be taken when interpreting empirical estimates of size effects on early life history stages, as they might not reflect reproductive patterns.

Nevertheless, the above studies illustrate compellingly the importance of carry-over effects in post-settlement selective processes. Although the sources of variations in offspring size, growth, and overall condition are subject to increasing research efforts, the underlying causality remains poorly understood.

*Parental effects and local adaptation*

Parental effects can be a major source of within-cohort variations in offspring phenotype (reviewed in Green 2008; Marshall et al. 2008). Potential pathways through which non-genetic parental effects can influence offspring phenotype and life histories include parental care, timing of spawning, female physiology, allocation of nutritional reserves, and egg composition (McCormick 1999; Green & McCormick 2005; Green 2008; Donelson et al. 2009). The nature of environmental conditions experienced by parents can thus codetermine the condition and fate of their progeny in two ways. Firstly, where the physiological condition of parents is affected (positively or negatively), parent quality may translate directly to progeny performance through resource and hormonal allocations (McCormick 2006; Gagliano & McCormick 2007). In general, larger parents in good condition produce larger, faster growing offspring with higher chances of survival over
successive life stages (Berkeley et al. 2004; Green 2008; Donelson et al. 2009). In contrast, adult exposure to unfavorable conditions can decrease offspring performance up to the point of recruitment and beyond. Dupont et al. (2012) for example recently demonstrated how adverse effects of adult exposure to elevated $p$CO$_2$ in sea urchins might translate into reduced settlement success of their offspring.

Secondly, adaptations to local environmental conditions can be transferred via parental effects, resulting in offspring that are well-adapted to the natal environment, as demonstrated in plant seed dispersal in terrestrial systems (Galloway 2005; Galloway & Etterson 2007). Marshall (2008) showed how such maternally induced phenotypic adaptations to local environmental conditions might come at the cost of less competitive performance if the offspring is exposed to non-natal environments in a marine bryozoan. Such phenotype-environment mismatches (sensu DeWitt et al. 1998) upon dispersal into new habitats might have strong influences on patterns of connectivity in the sea by acting as biological barrier to reproductive connectivity (reviewed in Marshall et al. 2010).

There is ample evidence for localized phenotypic adaptations in marine invertebrates, where common garden experiments revealed superior performance of local versus foreign individuals in the same selective environments (reviewed in Sanford & Kelly 2011). Yet disentangling genetic effects from those of phenotypic plasticity through parental effects remains challenging (Kawecki & Ebert 2004; Sanford & Kelly 2011).

Intriguingly, an individual’s dispersal behaviour might partially be controlled through maternal effects. As has been demonstrated in passerine birds, mothers can mediate their offspring’s dispersal propensity through differential transfer of maternal yolk androgens
(Tschirren et al. 2007). Similarly, environmentally induced maternal effects mediated seed dispersal distances in a flowering plant (Jacobs & Lesmeister 2012). Such adaptive parental control over dispersal strategies might be particularly useful in situations where local conditions (e.g. habitat availability) cannot readily be gauged by offspring, as is probably the case for most marine organisms with pelagic dispersal stages. There is sufficient evidence to show that maternal effects in brooding fish can have a substantial influence on offspring phenotype (e.g. McCormick 1999; Green & McCormick 2005; McCormick 2006). Yet it remains to be seen whether adaptive maternal provisioning in marine organisms can mediate dispersal of offspring according to environmental conditions.

**Dispersal environment**

Another non-genetic influence on larval and/or recruit phenotype is exerted directly via the environmental conditions experienced by the individual during different life history stages. Especially in benthic brooders, the egg stage will be subject to natal conditions until hatching occurs. Embryos might hence be ‘imprinted’ to the environment of their natal site (i.e. the local temperature, chemical, hydrodynamic, and turbidity regime). Upon hatching (or starting with the egg stage in broadcast spawners), larval/recruit phenotype might be influenced by the conditions encountered in the pelagic environment. Such ‘matrix effects’ (sensu Shima & Swearer 2009) include food availability (e.g. Meekan et al. 2003; Phillips 2004; Donelson et al. 2009), variations in temperature (e.g. Meekan et al. 2003; Sponaugle et al. 2006; Grorud-Colvert & Sponaugle 2011; Rankin &
Sponaugle 2011), solar radiation, wave height, wind regime, and rainfall (Bergenius et al. 2005; Macpherson & Raventos 2005). Recent research indicates that the pelagic dispersal matrix indeed has a significant influence on patterns of dispersal and settlement success (Shima et al. 2010). Using information stored in the otoliths of the common triplefin (Forsterygion lapillum), Shima & Swearer (2009; 2010) were able to link higher recruit quality (a composite measure of growth-related traits) to retention in local waters compared to long-distance dispersal. While these particular results were likely produced by strong heterogeneity between dispersal habitats (nutrient rich vs. nutrient poor), the studies strikingly establish the potential importance of different dispersal pathways for recruit condition. Local recruitment might hence not only be influenced by the interplay of hydrodynamic forcing with larval behaviour (Leis 2010), but also by individual dispersal histories (‘origin’ plus ‘dispersal pathways’).

Whether larval/recruit quality is a function of the source population (parental effects, imprinting), dispersal history (matrix effects), or a combination thereof will have vast implications for patterns of metapopulation connectivity and our ability to accurately predict them (Shima & Swearer 2009). Yet few studies have so far addressed this issue and the ones that did yield somewhat conflicting results (parental effects: Donelson et al. 2009; pelagic conditions: Shima & Swearer 2009). Contrasting outcomes are to be expected however, where inter- and even intraspecific variability can be high and methodologies diverge. Monro et al. (2010) demonstrated that selection for offspring size in bryozoans varies greatly between laboratory and field conditions and even between cohorts in the latter. Care should hence be taken when interpreting observed patterns under controlled lab conditions, as they may not reflect realized field conditions. So far
the relative importance of parental effects versus legacy effects (dispersal environment) on recruit quality remains largely unknown.

**Dispersal costs**

In addition to the potential impact of the dispersal environment on larval/recruit quality, the process of dispersal itself can bear a cost to individual condition (Bonte et al. 2012). Costs of dispersal can either be immediate (i.e. larval mortality), or deferred (i.e. occurring after settlement). Deferred costs can again be separated into direct and indirect deferred costs. **Direct costs:** Especially for non-feeding larvae, long-distance dispersal may be energetically costly, a price that is paid either right before (reduced effectiveness in habitat selection), and/or after colonization (reduced physiological condition) (Pechenik 2006; Marshall et al. 2010; Burgess et al. 2012). **Indirect costs:** As discussed above, adaptations to environmental conditions of the natal site might produce phenotype-environment mismatches where immigrants arrive at habitats to which they are not suited (Nosil et al. 2005; Marshall et al. 2010).

Immediate as well as deferred dispersal costs are believed to scale with increasing dispersal distance and/or time (Rousset & Gandon 2002; Stamps et al. 2005; Bonte et al. 2012). While there is evidence for selection against immigrants with longer distance dispersal from the terrestrial literature (e.g. Baker & Rao 2004; Matter 2006; Johnson et al. 2009), little is known of similar effects in the marine environment. Burgess et al. (2012) recently demonstrated that prolonged dispersal durations in a marine bryozoan (*Bugula neritina*) might result in poor recruit performance due to higher likelihoods of
colonizing low-quality habitat. Such direct deferred dispersal costs might be common in marine organisms and might regularly result in non-linear mismatches between potential (dispersal) and realized (post-settlement survival) connectivity (Burgess et al. 2012).

In theory, dispersal costs should favor individuals that thrive at the short end of the dispersal kernel (Nosil et al. 2005; Marshall et al. 2010; Bonte et al. 2012). While recent empirical evidence for short distance dispersal in coral reef fishes is supporting this hypothesis (e.g. Almany et al. 2007; Buston et al. 2011, Almany et al. 2013), the causal ecological processes underlying the observed patterns as well as the large variations in dispersal distances at the long end of the dispersal kernel remain largely unexplored.

Phenotypic dispersal predictors – evidence from terrestrial systems

Alternatively, long distance dispersal might be favored by selection where it is necessary to avoid (1) inbreeding, (2) kin competition and/or (3) environmental stochasticity (reviewed in Clobert et al. 2001). These mechanisms are in theory balanced with the costs of dispersal. The fitness of a dispersing organism would thus be a function of individual condition/phenotype and related dispersal costs, leading to variable levels of subsequent survival and/or reproductive success (Bernard & McCauley 2001). In other words, dispersal distance and reproductive success might be a result of the complex interactions between an individual’s phenotype and the environment it experiences (Clobert et al. 2009; Tarwater & Beissinger 2012). Depending on phenotypic predispositions, individuals experience and react to the same environment differently and
these interactions might influence which individuals disperse, how far they disperse and which return to the natal site (Clobert et al. 2009).

In principle it should hence be possible to identify specific morphological, physiological, or behavioural traits that are associated with dispersal and/or natal homing (Ronce & Clobert 2012). There is some evidence that dispersal strategies (i.e. distances) are related to certain intra-specific morphological predictors in different terrestrial organisms, such as wing span in birds (Dawideit et al. 2009), leg morphology in salamanders (Lowe & McPeek 2011), or size in small mammals (O’Riain et al. 1996; Selonen et al. 2012). Similarly, behavioural characteristics might predispose individuals to either leave or stay. ‘Boldness’ in particular is often shown to relate positively to dispersal (Cote et al. 2010). In coral reef fishes, larval behaviour is believed to be an important factor mediating stochastic hydrodynamic advection (Kingsford et al. 2002; Paris et al. 2007). Small differences in vertical migration behaviour, for instance, can have strong effects on dispersal distance and direction (Sundelöf & Jonsson 2012).

Intrinsic behavioural dispersal propensities like boldness might hence strongly influence dispersal pathways. At the same time, a lack of boldness (among other traits) was shown to result in increased post-settlement mortality in a damsel fish (Pomacentrus wardi) (Fuiman et al. 2010). Variations in traits such as individual boldness might hence influence propensities to dispersal as well as post-settlement selection, and might thereby present one selective mechanism balancing the costs and benefits of dispersal in marine organisms.

The accurate assessment, interpretation and extrapolation of dispersal phenotypes (dispersotypes) is hampered by the often strong covariation between different traits and
the intraspecific variability thereof depending on internal and external conditions (Ronce & Clobert 2012). Nevertheless, the identification of specific dispersal predictors would be useful in the (admittedly simplified) forecast of connectivity patterns. This is true especially in marine systems where dispersal itself is exceedingly difficult to measure. As to this moment, however, no specific phenotypic predictors of dispersal (e.g., larval fin aspect ratios, size, specific behaviours) have been identified and validated in any marine organisms.

Dispersal as a function of habitat fragmentation

As discussed above, larger geographic distances between patches theoretically increase the cost of dispersal, the correlation of which will depend on inter- and/or intra-specific dispersal potentials (Clobert et al. 2009; Bonte et al. 2012). The propensity to disperse away from the natal habitat was shown to decrease with increasing habitat fragmentation/isolation in terrestrial insects (Bowler & Benton 2009; Ahlroth et al. 2010), and such propensities might be genetically inherited (Ahlroth et al. 2010), maternally induced (Tschirren et al. 2007), and/or a function of habitat cues (Bowler & Benton 2011). While as yet mostly hypothetical, the propensity of marine larvae to return to the natal patch rather than disperse might accordingly evolve as a function of increases in habitat patchiness (Baskett et al. 2007). Probing of the water column by settlement stage larvae might accordingly trigger local retention, if habitat cues are distant, or dispersal, if foreign habitats appear to be closer. Such ‘decisions’ about dispersal – whether pre-programmed or environmentally induced – will undoubtedly be more
prominent in organisms with active larvae and among those, more pronounced in brooding species.

The importance of larval behaviour for patterns of connectivity has long been recognized (Leis & Stobutzki 1997; Atema et al. 2002; Kingsford et al. 2002; Leis 2006) and coral reef fish larvae in particular were repeatedly shown to exhibit oriented swimming behaviour sophisticated enough to influence dispersal trajectories (reviewed in Leis et al. 2011). Recent empirical studies identified variable patterns of dispersal between actively orientating (fish) and relatively passive (invertebrate) larvae. In invertebrate larvae, mean dispersal distances might be longer (López-Duarte et al. 2012) and patterns largely explained by hydrodynamic forcing (Gilg & Hilbish 2003). In contrast, fish larvae were shown to actively retain themselves in local waters (Sponaugle et al. 2002; Warner & Cowen 2002; Cowen et al. 2006) and patterns of inter-patch connectivity on restricted spatial scales might be largely independent of prevailing circulation patterns (Saenz-Agudelo et al. 2012). Moreover, connectivity patterns in reef fishes were shown to be remarkably stable over time (Berumen et al. 2012; Saenz-Agudelo et al. 2012), while invertebrate dispersal seems to vary significantly according to seasonal changes in current regimes (Carson et al. 2011; López-Duarte et al. 2012).

Behavioural predispositions might thus be important predictors of dispersal patterns in reef fishes. The evolution of ‘dispersotypes’ in marine systems clearly requires more dedicated research and future studies should integrate measures of phenotypic traits into empirical estimates of larval dispersal.
Origin effects

In the light of recently developing paradigms, we highlight the importance of linked life histories on individual phenotypes and the potential consequences of variations in the latter to patterns of larval dispersal and post-settlement selection. To conceptualize the interplay of the presented ideas (i.e. parental effects, local adaptation, dispersal history, individual phenotype, post-settlement selection, dispersotype, habitat fragmentation) into a comprehensive theory, we propose the concept of origin effects (Fig. 1).

**Figure 1:** Schematic showing the different processes that constitute origin effects: the natal environment might act upon embryonic and larval condition either directly (imprinting) or via parental effects; conditions in the pelagic environment might influence larval condition directly through matrix effects; carry-over, latent, and/or legacy effects from the larval/embryonic stage might act on recruit condition; the geographic distance between natal and settlement habitat will increase the necessary dispersal distance and might thereby increase direct deferred dispersal costs (DDDC), habitat differences between the natal and the settlement site might scale with distance and entail indirect deferred dispersal costs (IDDC) through phenotype-environment mismatches; the dispersotype might be shaped by a combination of environmental conditions, maternal conditioning, and/or geographic distance, reciprocally the dispersotype might determine dispersal distance.
The label ‘origin’ effects is not meant to necessarily emphasize the importance of the larval source itself, but rather signifies the combined effects (synergistic or antagonistic) of the different stations from the natal patch over dispersal and settlement through to reproduction. We define origin effects as the collective influences of natal origin and dispersal environment on recruit condition and/or survival in the context of habitat arrangement. In other words, origin effects conceptualize dispersal costs as a function of intraspecific variations in phenotype in conjunction with quantitative and qualitative aspects of the habitat matrix. To discuss and illustrate the concept, we will mostly focus on coral reef fishes, simply because recent research progress in the involved fields largely stems from these systems.

Semantics of connectivity research

Over the past few decades there has been a long-standing debate of whether metapopulation dynamics of reef fishes are mainly regulated by local self-seeding or connectivity through wide range dispersal between subpopulations (e.g. Sale 1991; Mora & Sale 2002; Warner & Cowen 2002). Recent evidence from empirical and modeling studies is establishing a picture of high self-recruitment (e.g. Berumen et al. 2012) and dispersal kernels describing rapidly declining probability functions with increasing distance (e.g. Buston et al. 2011; Horne et al. 2011, Almany et al. 2013). The tail of these kernels would be characterized by sporadic and irregular long distance dispersal, often sufficient to maintain genetic homogeneity between distant demes (Hedgecock et al. 2007). While we are thus starting to realistically describe patterns of larval dispersal,
there remains a lack of understanding about the origin and influence of post-settlement selection on patterns of realized connectivity (Pineda et al. 2007; Allen & Marshall 2010; Burgess & Marshall 2011; Burgess et al. 2012).

While the discussion about open versus closed populations is approaching a common denominator, integral issues related to semantic misinterpretations might pose an unnecessary hurdle in theoretical advancements. In this context, we would like to reemphasize the difference between self-recruitment and local retention in marine connectivity. Self-recruitment (SR) denotes the fraction of larvae that return to the natal patch, compared to the fraction that enters the population from (mostly unknown) external sources. Local retention (LR) on the other hand describes the fraction of larvae that return to the natal patch compared to the total number of larvae produced from that same patch. While SR can be estimated empirically using tagging (e.g. Almany et al. 2007) or parentage techniques (e.g. Planes et al. 2009), LR remains difficult to measure due to high levels of advection and larval mortality. Importantly, the two concepts are not necessarily linked and measures of SR do not automatically reflect levels of larvae returning to the natal reef (Fig. 2). Instead, LR might be decoupled from SR where ‘dilution’ from nearby surrounding patches increases the SR denominator, thereby reducing the estimates of SR although LR might be high (Pinsky et al. 2010; 2012). Alternatively, in more isolated patches, estimates of SR will increase where due to increased patch distance levels of incoming larvae from external sources are low (Pinsky et al. 2012). Levels of LR might hence be high at both ends of the spectrum of habitat fragmentation, or potentially even higher at clumped patches due to ‘sticky water effects’
(Andutta et al. 2012), while yielding different levels of SR (e.g. Jones et al. 2005; Saenz-Agudelo et al. 2011).

**Figure 2:** Potential scenarios where local retention and self-recruitment might be decoupled. (a) High retention at patch A due to the sticky water effect in dense habitat aggregations (Andutta et al. 2010) is masked by dilution effects through increased colonization from nearby patches B and C. (b) Relatively low retention in a more isolated patch A, but little input from distant external sources B and C (concentration effect). (c) Concentration effect at patch A due to selective mortality where dispersal costs result in higher mortality of external recruits from B and C compared to local recruits.
The impact of post-settlement selection

Patterns of SR and connectivity measured by most field methods have the added complication that they may reflect patterns arising after dispersal due to post-settlement processes. As discussed above, dispersal costs may indirectly favor natal homing and/or short distance dispersal in two ways: (1) The dispersal process itself can be costly: Increases in dispersal distance are predicted to result in reduced survival and/or fitness of recruits because of the energetic cost of dispersal (Bonte et al. 2012). (2) Phenotype-environment mismatches: Locally adapted phenotypes might result in a competitive advantage of local recruits versus individuals from source populations that might bear a different phenotypic signature (Nosil et al. 2005; Marshall et al. 2010). Both points should theoretically correlate with patch distance and/or isolation. Estimated post-settlement patterns of SR might hence relate to habitat fragmentation in that local recruits might increasingly be favored over external recruits with increasing distances that need to be covered in order to successfully disperse. Such selective post-settlement mortality would similarly affect the observed patterns of SR, as would a lack of recruits from external sources (Fig. 2).

Recent research provides compelling indications for direct (Burgess et al. 2012) and indirect (Marshall 2008; Marshall et al. 2010) deferred dispersal costs and emphasizes the importance of dispersal history on post-settlement selection (Shima & Swearer 2009; 2010; Shima et al. 2010; Smith & Shima 2011). There is mounting evidence for the potential cumulative effects of natal origin, dispersal duration (distance) and environmental heterogeneity (‘legacy effects’) to skew reproductive dispersal kernels towards the origin (Marshall et al. 2008; Shima & Swearer 2009; Allen & Marshall 2010;
Recent empirical measures of short distance dispersal, taken together with the discussed theoretical and direct indications for post-settlement selection against immigrants (admittedly partially derived from invertebrates) leads to the question whether dispersal in coral reef fish larvae is merely an accident. This notion is supported by behavioral studies that demonstrated active orientation towards the natal reef in settlement stage larvae (Gerlach et al. 2007) and biophysical mechanisms of local retention (Warner & Cowen 2002; Cowen et al. 2006). Intraspecific variation in larval quality might hence produce varying fractions of actively retained and passively dispersed larvae in each cohort (see Armsworth et al. 2001). While larger and faster growing (high quality) larvae might use their better-developed behavioural mechanisms (sensory orientation and swimming) to be retained near the natal site, lower-quality individuals might be flushed out by prevailing currents (Fig. 3). Recent empirical support for this hypothesis stems from a study by Beldade et al. (2012) employing parentage analysis to estimate self-recruitment in an anemonefish (A. chrysopterus) around Moorea, French Polynesia. Interestingly, they found that larger females contributed disproportionally to local replenishment, producing more natal recruits than smaller conspecifics (Beldade et al. 2012). The authors speculate that maternal effects might act upon growth-related larval traits leading to higher survival in offspring of larger mothers. As an alternative explanation, what they found might also indicate a higher proportion of maternally induced high-quality recruits returning to the natal reef, while lower quality individuals (i.e. produced by smaller mothers) more or less passively disperse. Whether
any of these speculations holds true remains to be seen. In either case, the study raises an interesting point and exemplifies a more experimental way of employing parentage analysis.

Figure 3: Schematic indication of larval dispersal: Spawning at patch (population) A results in a cloud of pelagic larvae of different quality (here: high or low) through genetic and/or parental effects. From this larval pool, different fractions are retained at the natal patch, disperse to external patches, or are advected away from any suitable habitat (equaling death). While a large fraction of high quality larvae might actively retain themselves at the natal patch, higher fractions of low quality larvae (less capable of active navigation) are advected towards external patches. The further the patch, the higher the fraction of low quality individuals through: (a) retention of high quality larvae at closer sites, and (b) the cost of dispersal. In this illustration, more larvae originating at patch A will hence successfully settle at patch A than at patch B and at patch B than at patch C. Similarly, high or low larval quality could be replaced with dispersotypes (local or long distance).

Conclusions

A recently (re)-emphasized theme in the field of marine metapopulation connectivity describes the influence of habitat quantity (geography) and quality (heterogeneity) on
patterns of larval dispersal and post-settlement selection. The complex interaction between parental and environmental effects on larval/recruit condition and the influence of carry-over effects between life history stages is receiving increasing emphasis in the literature, while at the same time, empirical estimates and model simulations highlight the importance of larval/recruit quality on patterns of realized connectivity.

Empirical estimates of larval dispersal allow for the first time to incorporate actual patterns of propagule exchange and retention into models of marine population connectivity, enabling field-validations of theoretical dispersal models. However, while such model validations are sorely needed (Werner et al. 2007), it will be necessary to disentangle between patterns of larval dispersal (model simulations) and realized connectivity \((post\ selection)\). While parentage analysis and population genetic studies differ in the temporal demographic scales at which they operate, they both have one thing in common: sampling is usually conducted at relatively late stages of ontogeny, i.e., past the bottleneck of post-settlement selection. Care should hence be taken when evaluating model performance based on either genetic approach.

Clearly not all individuals arriving at a recruitment habitat will have equal chances of surviving through settlement and to reproduction. Instead individual recruit performance will depend on dispersal histories and carry-over effects leading to intra-specifically variable phenotypes. Future biophysical dispersal models should hence seek to incorporate the influence of origin effects on post-settlement performance.

To supply modelers with more solid estimates of the influence of origin effects, we need to develop more innovative ways of tackling the problem. Despite a growing
interest in this field of research, studies pertaining to the interplay between larval histories, phenotypic traits and selective mortality in the sea are facing some critical methodological challenges, particularly regarding non-sessile organisms (Johnson et al. 2012). For once, high rates of mobility and natural mortality preclude continuous measures of traits in specific individuals over time. Moreover, the identification of dispersal past as well as most traits of interest requires the lethal sampling of individuals. Many studies hence employ a cross-sectional approach during which specific traits are measured in different individuals before and after a selective event. Distributions of traits are then compared between pre-selection (e.g. larvae) and post-selection samples (e.g. recruits) (Johnson et al. 2012). Many of the mentioned studies in this paper reconstruct early life history traits in post-settlement individuals via incremental depositions in hard structures such as otoliths in fish (e.g. Hamilton et al. 2008), or shells in invertebrates (e.g. Carson et al. 2011). Analysis of isotopic composition of these increments may assist in the identification of individual larval sources and/or dispersal environments (e.g. Shima & Swearer 2009, but see Berumen et al. 2010). While these kinds of analyses have proven to be invaluable for the reconstruction of larval histories in the light of selective survival, we argue that the power of such approaches can be raised significantly if combined with empirical techniques.

Parentage analysis has so far been used in a purely descriptive way to measure self-recruitment and to a lesser degree connectivity between adjacent reefs (e.g. Jones et al. 2005; Planes et al. 2009; Saenz-Agudelo et al 2011; 2012). While the technique has already generated important insights into connectivity patterns, it has not yet been employed to its full potential. Beldade et al. (2012) pointed towards a more experimental
way of using parentage by relating levels of self-recruitment to maternal size. Similarly, if employed in a more explicit background, the power to identify natal and external recruits could be used to estimate a whole range of components of origin effects.

While the principles that constitute the theory of origin effects are receiving increasing attention in the literature, the difficulty of measuring any aspect of the concept in the field have mostly hampered progress beyond single features over limited time scales. On the basis of recent findings, we have shown the influence of origin effects on patterns of realized connectivity. We hence advocate a comprehensive empirical approach to study the different processes acting in concert to create origin effects.
References


CONCLUSIONS

The processes shaping patterns of population connectivity are gaining increasing attention in the marine ecological literature, not least because of the urgent need to incorporate larval dispersal into the conservation management of threatened coastal ecosystems (Botsford et al. 2009). Depending on the observational temporal and spatial resolution, we differentiate between evolutionary and ecological (demographic) connectivity (Chapter 1). Most connectivity studies to date have approached the subject from one angle on either end of the spectrum. Interspecific patterns of evolutionary and demographic connectivity, however, are often less consistent than expected. Coral trout (Plectropomus maculatus) and stripy snapper (Lutjanus carponotatus) for instance were shown to lack significant genetic structure across the Great Barrier Reef, indicating long distance dispersal over 1000s km (Evans et al. 2010). On the other hand, empirical estimates of dispersal on smaller spatial scales revealed high percentages of localized recruitment in both species (Harrison et al. 2013). Similarly, the anemonefish Amphiprion clarkii exhibited little structure across the Great Barrier Reef (Doherty et al. 1995) on evolutionary scales. Recent empirical studies on the contrary have produced high estimates of self-recruitment (20-60%) in closely related species, Amphiprion polymnus and Amphiprion percula on demographic scales (Jones et al. 2005; Almany et al. 2007; Planes et al. 2009; Saenz-Agudelo et al. 2011; Berumen et al. 2012). The combination of direct and indirect approaches will provide a more holistic picture of connectivity over a range of spatial and temporal scales (Burbrink 2010; van de Meer et al. 2013).
In this thesis we addressed larval connectivity from two different angels. First we implemented a large-scale population genetic approach to resolve patterns of evolutionary connectivity across the entire Red Sea basin (Chapter 4). Secondly, we employed direct genetic methods in combination with a biophysical model to assess patterns of demographic connectivity at more restricted spatial and temporal scales (Chapter 5). Both approaches created a picture of extensive gene flow along the Saudi Arabian Red Sea coast with little localized recruitment.

The combination of large-scale population genetic patterns with environmental data (chlorophyll-a concentrations and Euclidian geographic distance) resulted in a highly informative model of isolation by environment \( R^2 = 0.92 \), reinforcing the concept of landscape genetics in marine systems (Selkoe et al. 2008; Liggins et al. 2013). This study demonstrates how informative aspects of the seascape can be employed to predict and interpret large-scale genetic differentiation. We support growing calls for the inclusion of habitat characteristics into future studies of genetic structure as well as into the parameters of dispersal models.

The combination of direct genetic methods with a biophysical model on contemporary connectivity scales yielded the first empirical evidence of a largely open population of coral reef fish. This finding is particularly noteworthy when compared with previous studies, most of which have found high levels of self-recruitment in coral reef fishes, regardless of the empirical approach (e.g. modeling: James et al. 2002 [55%]; geochemistry: Patterson et al. 2005 [75%]; barium tagging: Almany et al. 2007 [60%]; parentage analysis: Planes et al. 2009 [42%]). We thus challenge the emerging paradigm of universally high levels of local recruitment in coral reef fish populations (e.g. Jones et
al. 1999; Swearer et al. 2002; Paris & Cowen 2004; Jones et al. 2005; Gerlach et al. 2007; Buston et al. 2011; Almany et al. 2013) and urge for caution when assuming self-replenishment in marine reserve design.

The regulatory drivers behind the observed patterns of larval dispersal remain enigmatic. While physical forcing is clearly a major factor, different aspects of larval biology might be equally important (Shima et al. 2010; Leis et al. 2011; Burgess et al. 2012). At least at smaller spatial scales, specific groups of larvae may exhibit individual tendencies towards dispersal or local retention. These “dispersotypes” (Ronce & Clobert 2012) may be regulated by “origin effects” (Chapter 6), which describe the influence of dispersal history (i.e. natal origin plus dispersal environment) on realized connectivity.

Overall, the combined approaches employed in this thesis revealed significant new insights into the patterns of larval connectivity in coral reef fishes. Over the past decades, the Red Sea has notoriously been understudied compared to other ocean basins (Spaet et al. 2012; Berumen et al. 2013). Here we present vital data on patterns of gene flow and larval dispersal along the Saudi Arabian Red Sea coast that will be useful for the effective implementation of spatial management strategies in this basin and beyond.
References:


APPENDIX 1

List of publications

Publications arising from this thesis:


This paper resulted from Chapter 3 and describes the development of the genetic markers used in Chapters 4 and 5. Contribution: G.N. collected the field samples, extracted DNA, and conducted laboratory work and data analyses.

2. **Nanninga GB**, Berumen ML *(in review)* Origin effects in marine population connectivity. *Fish and Fisheries*

This synthesis/concept paper is presented in Chapter 6 and is in review for *Fish and Fisheries* since June 2013. Contribution: G.N. devised and wrote the manuscript.


This paper is presented in Chapter 4. The manuscript is in preparation and will be submitted to *Molecular Ecology* by the end of September 2013. Contribution: G.N. designed the study, collected samples, conducted laboratory and statistical analyses, and wrote the paper.

This paper is presented in Chapter 5. The manuscript is in preparation and will be submitted to non-specified peer-reviewed journal by the end of September 2013. 

*Contribution*: G.N. designed the study, collected samples, conducted laboratory and statistical analyses, and wrote the paper.


This paper is not presented in the thesis. The manuscript is based on data from Chapter 5 and will describe the increase in false positive parent-offspring assignments when the number of genetic markers is reduced. By the time of finishing this thesis, data analysis was still in progress. The manuscript will be submitted to *Molecular Ecology* by the end of November 2013. 

*Contribution*: G.N. designed the study, conducted laboratory and statistical analyses, and wrote the paper.


This paper is not presented in the thesis. The manuscript will be based on data from Chapter 4 and will describe a spatially explicit genetic modeling framework informed by ocean circulation. By the time of finishing this thesis, model development was still in progress. The manuscript will be submitted to a non-specified peer-reviewed journal by the end of January 2014. 

*Contribution*: G.N. supplied the large-scale genetic data that forms the basis of this approach (Chapter 4) and will be involved in writing the paper.
Additional publications:


This paper describes the development of 27 genetic markers for *Dascyllus marginatus.*

*Contribution:* G.N. assisted in the collected of field samples and laboratory analysis.

8. Nanninga GB, Manica A, Berumen ML (*in preparation*) Social hierarchies are shaped by habitat: intraspecific variations in size ratios among host anemones in the clownfish *Amphiprion bicinctus.* *Ecology*

This paper is based on size data collected during fieldwork on Chapter 4 and will describe the difference in size hierarchies among species of host anemones, as well as distributional patterns of *A. bicinctus* along the Red Sea. By the time of finishing this thesis, data analysis was still in progress. The manuscript will be submitted to *Ecology* by the end of December 2013. *Contribution:* G.N. devised the study, collected data, is conducting statistical analyses, and will write the paper.
APPENDIX 2

Non-thesis related fieldwork participation during PhD

2 - 16 Apr 2013: Liveboard - Papua New Guinea: field collections of *Centrophyge bicolor* and *Amphiprion percula* for large scale empirical connectivity project. PI: Dr. G. Jones

20 - 23 Mar 2013: PI on research cruise - Red Sea: testing and deployment of the In Situ Ichthyoplankton Imaging System (ISIIS)

24 Sep - 5 Oct 2012: Liveboard - Djibouti: field collections of various species of reef fishes and invertebrates. PI: Dr. J. DiBattista

1 - 24 June 2012: Liveboard - Phoenix Islands, Kiribati: field collections of various reef fish species, satellite-tagging of *Manta birostris*. PI: Dr. S. Thorrold

27 Sep - 4 Oct 2011: *A. bicinctus* collections - Egypt: unsuccessful due to permit issues.

25 Mar - 29 Apr 2011: Research trip - Papua New Guinea: land-based collections of *Chaetodon vagabundus* and *A. percula*. PI: Dr. G. Almany

4 - 19 Jan 2011: Dive training - Cayman Islands: Closed Circuit Rebreather dive training leading up to certification.
APPENDIX 3

List of conference presentations during PhD

2013: 9th Indopacific Fish Conference, Naha, Japan (oral presentation)

2013: 37th Annual Larval Fish Conference, Miami, USA (oral presentation)

2012: 12th International Coral Reef Symposium, Cairns, Australia (oral presentation)

2012: ICES/PICES Conference for Early Career Scientists, Calvià, Spain (oral presentation)

2010: International Society for Reef Studies Symposium, Wageningen, Netherlands (poster presentation)