Coupled biophysical global ocean model and molecular genetic analyses identify multiple introductions of cryptogenic species

Michael N. Dawson†‡, Alex Sen Gupta§, and Matthew H. England§

1School of Biological, Earth, and Environmental Sciences, 2Centre for Environmental Modelling and Prediction, School of Mathematics, University of New South Wales, Sydney NSW 2052, Australia

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved July 5, 2005 (received for review May 7, 2005)

The anthropogenic introduction of exotic species is one of the greatest modern threats to marine biodiversity. Yet exotic species introductions remain difficult to predict and are easily misunderstood because knowledge of natural dispersal patterns, species diversity, and biogeography is often insufficient to distinguish between a broadly dispersed natural population and an exotic one. Here we compare a global molecular phylogeography of a representative marine meroplanktonic taxon, the moon jellyfish Aurelia, with natural dispersion patterns predicted by a global biophysical ocean model. Despite assumed high dispersal ability, the phylogeny reveals many cryptic species and predominantly regional structure with one notable exception: the globally distributed Aurelia sp. 1, which, molecular data suggest, may occasionally traverse the Pacific unaided. This possibility is refuted by the ocean model, which shows much more limited dispersion and patterns of distribution broadly consistent with modern biogeographic zones, thus identifying multiple introductions worldwide of this cryptogenic species. This approach also supports existing evidence that (i) the occurrence in Hawaii of Aurelia sp. 4 and other native Indo-West Pacific species with similar life histories is most likely due to anthropogenic translocation, and (ii) there may be a route for rare natural colonization of northeast North America by the European marine snail Littorina littorea, whose status as endemic or exotic is unclear.

biodiversity / biogeography / conservation

Until recently (1), marine species introductions were of limited concern because many marine plankton were assumed to have naturally broad, even global, distributions (2–4). However, as marine molecular genetics and physical oceanography have increasingly revealed biotic and physical discontinuities in an ecologically heterogeneous environment (4–6), anthropogenic introduction of nonindigenous species (NIS) has been recognized as a major threat to native biodiversity (7, 8) with the potential to alter ecosystems (9, 10), displace endemic species, and cost millions of dollars in damage and preventative control (11). Yet the threat is still poorly understood (7) and species introductions remain generally unpredictable (12) because knowledge of natural dispersal patterns (4, 6), species diversity (13), and biogeography (14) is often insufficient to identify NIS or their sources. Recent surveys classified 36–47% of NIS (15, 16) as "cosmopolitan" (i.e., distributed globally) (8, 15).

Molecular genetic techniques have been heralded as a powerful tool for taxonomic verification, differentiating cryptic taxa, identifying source populations and vectors, and assessing the extent and impacts of invasions (16, 17). However, this approach to NIS relies on the questionable assumption that introduced populations are characterized by massively reduced genetic diversity and no unique genotypes (17), which will generate estimates of gene flow dating only to the last few centuries. Ocean-going vessels typically visit numerous ports multiple times carrying sufficient numbers of planktonic, colonial, or aggregating fouling invertebrates for successful reproduction and invasion (18). This recurrent threat has the potential to increase genetic diversity in the invaded range, creating, with sampling error, patterns indicative of endemicity predating human transoceanic voyages. As a result, molecular genetic analyses alone may be insufficient to assess whether a geographic occurrence represents a species introduction. Here, we demonstrate that simulating natural dispersion over the time scales of human travel aids identification of NIS and, at the same time, elucidates taxonomy, patterns of marine biodiversity, and biogeography.

Materials and Methods

Molecular Analyses. Tissues from 78 Aurelia medusae were preserved in 70–90% ethanol or salt-saturated dimethyl sulfoxide (19). Mitochondrial cytochrome c oxidase subunit I (COI) was amplified by using primers LCO1490f (20) and HCO2198f (21) and PCR conditions described in ref. 22. Nuclear ribosomal DNA (nrDNA) was amplified from 14 of these medusae, representing each major COI clade, by using primers jJTTS1 and 28S-2R and PCR conditions described in ref. 22. COI amplicons were purified by using PCR clean-up columns (Qiagen, Valencia, CA), and rDNA amplicons were cloned by using TOPO TA (Invitrogen) and purified by using the Amersham Pharmacia Flexiprep kit. Purified amplicons were labeled with BigDye and sequenced on Applied Biosystems 377 automated sequencers according to Applied Biosystems protocols at the University of New South Wales’ Ramaciotti Centre. Electropherograms were checked, and misreads were corrected. Homologous sequences from prior studies (20, 22–25) were appended to each data set. COI sequences were aligned on the basis of the amino acid translations, nrDNA sequences were aligned by using several gap-opening/gap-extension weighting schemes in CLUSTALX (26), and alignments were corrected by eye. Positions with missing data were excluded from subsequent analyses. The total data set comprises 240 medusae from 77 sites worldwide of which 174 were sequenced for COI and 155 for nrDNA (see Table 2, which is published as supporting information on the PNAS web site). Because of the large number of specimens, the phylogenetic analyses reported here use a subset of sequences for which both COI and nrDNA data were available and represented major reciprocally monophyletic clades found in prior studies and/or preliminary maximum parsimony analyses of the entire data set.

© 2005 by The National Academy of Sciences of the USA
These clades were assigned species status (20, 22–25) if recovered in analyses of mitochondrial and nuclear markers and if separated by sequence differences comparable with or greater than those seen between traditionally recognized morphospecies Aurelia aurita and Aurelia labiata [i.e., \( \approx 10\% \) in COI and ITS1 (23)].

Gene trees were reconstructed by using maximum likelihood in PAUP*4.0b10 (27) using relevant models of molecular evolution (COI, GTR + G; nrDNA, TrNef + 1 + G) identified by MODELTEST (28). Robustness of the Aurelia sp. 1 sister taxon relationship identified by phylogenetic analyses was tested against all other possible pairings by using the Kishino–Hasegawa test in PAUP*4.0b10 employing RELL bootstrap (100,000 replicates) and a one-tailed test (29). Bootstrap analyses consisted of 1,000 realizations using the same maximum likelihood models and also unweighted maximum parsimony (10,000 realizations) including gapped positions in nrDNA. Additional COI data describing Aurelia sp. 1 were used to reconstruct an haplotype network by using unweighted maximum parsimony in PAUP*4.0b10. Haplotype and nucleotide diversity and population subdivision (\( \delta \sigma \)) were calculated in ARLEQUIN 2.0 (30).

Ocean Biophysical Modeling. To investigate the limits of natural dispersion of species of Aurelia over multidecennium time scales, we developed a global Lagrangian model incorporating representative life-history characteristics of Aurelia (see Fig. 5, which is published as supporting information on the PNAS web site). These life-history characteristics include a small, probably perennial, benthic polyp that reproducibly asexually to produce other benthic polyps and free-living planktonic medusae; the medusa reproduce sexually, producing a planula larva that is brooded for a short period by the female just released into the water column where it spends probably \(<1 \) week before settling on the benthos where it metamorphoses into a polyp. Thus, the medusa is the main dispersive phase; medusae usually live \(<6 \) months in nature, although medusae \( >1 \) year old have been recorded; they may live up to 2 years in captivity (31, 32).

The circulation model is driven by monthly advection fields derived from a 20-year integration (1979–1998) of the Parallel Ocean Climate Model (33–35). The current version simulates ocean circulation from 68°N to 75°S with average horizontal resolution of \( \approx 0.25^\circ \). Surface forcing consists of 20 years of daily varying momentum, heat, and freshwater fluxes derived from European Centre for Medium-Range Weather Forecasts reanalysis fields (35). The Parallel Ocean Climate Model accurately represents surface circulation patterns (34) and interior watermass pathways (36). We take the ocean to be in a steady-state seasonal cycle derived from the 20-year mean of the simulation, thus ignoring interannual variability such as that due to the El Niño Southern Oscillation and millennial scale fluctuations in the ocean’s thermohaline circulation. This simplification reflects the relative stability of the Earth’s climate during the current interglacial period (37). The sensitivity of the simulated advection pathways to this assumption of steady-state ocean circulation is addressed in Fig. 6, which is published as supporting information on the PNAS web site, with none of our conclusions altered with regard to medusa advection ranges.

Each experiment is based on the release of \( \approx 20,000 \) particles from known Aurelia sp. 1 zones of occurrence with particle positions \( x_0 \) at time \( t \) given by

\[
x_t = x_{t-1} + U(x_{t-1}) \Delta t + R_N \sqrt{2K_b \Delta t} + C_{\max},
\]

Here \( \Delta t \) represents the model time step (6 h), and \( U \) is the flow velocity interpolated to the position of the particle. Particles are forced to remain active by ignoring any component of \( U \) that would cause the particle to become grounded. Eddy advection is included by means of a constant horizontal diffusivity (\( K_h = 1,000 \text{ m}^2 \text{s}^{-1} \)) formulated as a random-walk term where \( R_N \) is a normally distributed random number. We tested the effects of halved, doubled, and 5-fold greater diffusivities, and none of our conclusions regarding natural Aurelia dispersal were altered (see Fig. 7, which is published as supporting information on the PNAS web site). Lower values, such as \( K_h = 100 \text{ m}^2 \text{s}^{-1} \) (used to investigate krill transport in the Southern Ocean) (38), only limit dispersal potential. Our experimental philosophy is to run with generous diffusivity so that the particles’ simulated range extent is an upper bound on natural dispersal.

To improve particle migration through “island hopping,” the Parallel Ocean Climate Model land mask is supplemented with a high-resolution bathymetry data set (see Fig. 8, which is published as supporting information on the PNAS web site). Ocean gridpoints that contain islands are assigned a fractional number indicating the percentage of the grid box containing land. During the simulations, for each successive re-release, particles are released from these “island” grid-boxes in proportion to the number of particles that have impacted on the region during the previous year, scaled by the land fraction.

Release of the Aurelia particles occurs over an entire year but is biased toward summer months, corresponding to normal patterns of abundance of Aurelia medusae (32). Each particle advects through the surface 75 m of the ocean for one “lifetime” (i.e., up to 365 days) of approximately the upper life span of Aurelia (31, 32) after which time it “dies.” For each model grid box, a cumulative total of the number of particles in that grid box per time step is calculated, providing a cumulative occurrence distribution (COD) representing the range of possible positions (and time spent at each location) at which a medusa may have released planulae larvae, potentially resulting in colonization of the area by the benthic polyp stage. Once all particles have exceeded their lifetimes, a new release of particles is initiated with new coastal release locations calculated as a function of the coastal COD values. This process is repeated, at the completion of each set of lifetimes, by using the newly generated COD.

To represent unresolved coastal flows, new release locations determined by the COD are recalculated, before each successive re-release, by applying an effective diffusion along the coastline as a result of estimated tidal currents (\( C_{\max} \)) (see Fig. 9, which is published as supporting information on the PNAS web site). An additional background along coastal flows (1.0 m s\(^{-1}\)) is added to account for any nontidal flows (e.g., coastal trapped waves, transient coastal currents, and shelf waves). The tidal current speeds are derived from a global inverse tidal model (39, 40). A temperature (\( T \))-dependent modification also is made before each new release, with the COD being reduced by a factor \( \sigma \) for temperatures above and below the temperature maximum (\( T_{\max} \)) and minimum (\( T_{\min} \)) of known Aurelia habitats, such that

\[
\sigma(x, y) = \frac{1}{2\Delta T(x, y)},
\]

where

\[
\Delta T(x, y) = T_{\min} - T(x, y), \quad \text{for } T < T_{\min}
\]

\[
\Delta T(x, y) = T(x, y) - T_{\max}, \quad \text{for } T > T_{\max}
\]

Values for \( T_{\max} \) and \( T_{\min} \) depend on the species of Aurelia under investigation (e.g., for sp. 1, \( T_{\max} = 24.5{ }^\circ \text{C} \) and \( T_{\min} = 14{ }^\circ \text{C} \)). The exponentially decreasing rate of survivorship with increasing temperature deviation (\( \Delta T \)) is a general approximation derived from data describing temperature-dependent mortality in the scyphozoan jellyfish Mastigias (41). Mortality rates in Mastigias increase slowly with increasing deviation above average temperatures to a point at which mortality rate begins to increase rapidly, a pattern that is broadly consistent with data on other marine invertebrates, such as corals (e.g., ref. 42) and gastropods (43).

Dawson et al.
Distributions reach an approximate steady state within a century of model simulation time. This state is expected as the life history of *Aurelia* limits the geographic extent over which the medusae can advect or mix, even considering possible routes of migration by means of island stepping stones and coastal-zone diffusion. As a result, only the Japan release experiment was integrated for a full 10,000 years; all other experiments were integrated for 1,000 years, well beyond the simulation’s quasi-steady state.

**Results and Discussion**

The global phylogeny of *Aurelia* reveals at least 16 phylogenetic, i.e., 13 cryptic, species (Fig. 1; see also Figs. 10 and 11, which are published as supporting information on the PNAS web site), most of which appear to be regionally restricted. Several species, however, have disjunct distributions, which may be a characteristic of introduced species. These species include *Aurelia* sp. 4, which occurs in the Western Pacific and in Pearl Harbor typical of introductions dating to the Second World War (15); *A. aurita*, which is endemic to the North Atlantic and disjunct in the Black Sea like the introduced ctenophore *Mnemiopsis* (9); and *Aurelia* sp. 8, which has a “Lessepsian” distribution, i.e., occurs on both sides of the Suez Canal, typical of other introduced species including the scyphozoan jellyfish *Rhopilema nomadica* (44). Most remarkable is *Aurelia* sp. 1 (a cryptic species of *A. aurita*), which occurs in major warm-temperate regions around the globe (see Figs. 1 and 3). The sister-taxon relationship between *Aurelia* sp. 1, *A. limbata*, and *Aurelia* sp. 10 (see Table 3, which is published as supporting information on the PNAS web site) indicates that *Aurelia* sp. 1 is endemic to the western North Pacific and, therefore, dispersed globally from Japan. The anomalously broad distribution of *Aurelia* sp. 1, however, is suggestive of anthropogenic introduction, because COI shows reduced molecular diversity in Australia and California compared to Japan (Fig. 2) and divergence times estimated by using a scyphozoan COI “molecular clock” do not exclude the modern day (Table 1). However, average estimated divergence times precede global shipping by thousands or tens of thousands of years (Table 1) suggesting rare, but natural, long-distance dispersal on evolutionary time scales, consistent with the hypothesis that *Aurelia* sp. 1 is naturally globally distributed (2, 24) and that 95% confidence intervals encompassing zero divergence times simply reflect ongoing gene flow.

Anthropogenic range expansions are distinguished by the fact that they exceed an organism’s natural dispersal ability. Our model translocation results (Figs. 3 and 4) demonstrate that *Aurelia* can occupy ranges of thousands of kilometers and presumably disperse over these geographic distances on evolutionary time scales, consistent with interpolated ranges of most phylogenetic species of *Aurelia* (Fig. 1). However, natural dispersal and mixing occur only within limited geographic regions. For example, in the North Pacific, *Aurelia* spread from Japan northwards and eastwards in the Kuroshio Current, as well as mixing locally in the Yellow Sea, the East China Sea, and the Sea of Japan (see Fig. 12, which is published as supporting information on the PNAS web site). Dispersal beyond this area is limited because the simulated 1-year
Table 1. Estimated divergence times between Pacific populations of Aurelia sp.1

<table>
<thead>
<tr>
<th>Populations</th>
<th>$d$</th>
<th>95% CI</th>
<th>$T$,†</th>
<th>95% CI,‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>California–Australia</td>
<td>0.036</td>
<td>±0.038</td>
<td>12,000</td>
<td>±13,000</td>
</tr>
<tr>
<td>California–Japan</td>
<td>0.431</td>
<td>±0.038</td>
<td>147,000</td>
<td>±13,000</td>
</tr>
<tr>
<td>Australia–Japan‡</td>
<td>0.202</td>
<td>±0.028</td>
<td>69,000</td>
<td>±10,000</td>
</tr>
</tbody>
</table>

Divergence time and 95% confidence interval (CI) estimated from net pairwise divergence (17) $T = d/2$ where $d = D_{xy} - (D_x + D_y)/2$ and $\lambda$ (point mutation rate per year) $= 1.47 \times 10^{-8}$ (ref. 20; see also refs. 46 and 47). $D_{xy}$ = average number of pairwise differences between regions; $D_x$, $D_y$ = average number of pairwise differences within regions.

†Divergence time $T$ is measured to the nearest millennium.
‡From ref. 20.

The discrepancy between the implications of molecular and ocean modeling analyses highlights two genetic effects with
important consequences for our understanding of global NIS. First, estimates of local diversity in NIS populations are inflated by multiple introductions, thus masking the expected low-diversity genetic signature of introduced species. Second, estimates of the timing or duration of gene flow can be inflated by multiple introductions, especially if dispersal by means of human vectors favors differential establishment of alleles rare in the natural range. Both erode strong genetic signals expected in NIS, giving the false appearance of a naturally dispersed endemic population.

The ocean model also can address other anthropogenic introductions. For example, *Aurelia* sp. 4 is endemic to eastern Borneo and Palau and also occurs in Hawaii (Fig. 1). The ocean model shows that *Aurelia*-like particles released off Borneo are advected eastwards to Palau by an eddy in the Celebes Sea and via the Equatorial Counter Current but that none disperse to Hawaii even on multicentury time scales (Fig. 4). This result is robust even under the most extreme and unrealistic dispersal scenario of a perpetual El Niño circulation field, demonstrating that for a species with the dispersal characteristics of *Aurelia* sp. 4, there is no available ocean pathway that naturally connects these zones of occurrence on these time scales. The range-limit indicated in the model (Fig. 4) is consistent with preliminary collections that found *Aurelia* sp. 11 (not sp. 4) in the Marshall Islands and, more generally, with Indo-West Pacific biogeography wherein many species that occur in the Indo-West Pacific center of marine biodiversity do not penetrate into the central Pacific. For example, there are 2,500 species of shorefishes in the Philippines, 1,300 in Palau, 850 in the Marshall Islands, and only 550 in Hawaii (49, 50); similar patterns are evident in scyphozoan jellyfishes (51, 52), corals (53), and many other taxa. Concomitantly, historical records reveal that *Aurelia* is not native to Hawaii. Despite surveys of scyphozoans in the early 1900s (54, 55), the genus *Aurelia* was first reported in Hawaii in Pearl Harbor in 1953 (56) after considerable WWII naval traffic with the western Pacific (J. Carlton, personal communication). This reciprocity of historical, biogeographic, molecular, and model data helps validate the model’s solution, which now provides additional support to the existing body of evidence (e.g., refs. 15 and 57) that a substantial number of endemic Indo-West Pacific species that also occur in Hawaii but have limited natural dispersal ability, including other scyphozoans with life histories like *Aurelia* sp. 4 (58–60), are most probably introduced.

Like evidence of disjunction, evidence of connectivity in the ocean model is also important for interpreting data on species introductions. The model identifies a possible rare route of natural dispersal across the North Atlantic, via the northern limb of the subpolar gyre (Fig. 3), potentially supporting recent genetic evidence that *L. littorea* populations in northeast North America are endemic resulting from natural range expansion predating Viking expeditions (17). The projected pathway is sensitive to the choice of temperature dependence and model boundary conditions, but the route identified by the model should, if anything, be stronger for the boreal *L. littorea* than those shown in Fig. 3 for the warm-temperate *Aurelia*. This rare dispersal route is consistent with evidence of different varieties of *A. aurita* in Europe and North America (2, 23) and of many east–west amphip-Atlantic species, particularly boreal and Arctic-boreal molluscs and fishes (61). However, the conclusion of natural range expansion based on genetic data (17) is not indisputable, and it remains to be demonstrated that observed patterns and levels of genetic diversity are explained better by rare dispersal of *L. littorea* than by multiple introductions.

The patterns of connectivity and disjunction identified by the genetic data and ocean trajectory model are also broadly consistent with other previously recognized biogeographic patterns (61). For example, the highest particle densities in the Australian simulations are concordant with the Flindersian Province, the eastern boundary of which is marked by an abrupt decrease in transport around Cape Howe (the southeastern tip of mainland Australia), although some propagules do enter the eastern transition zone, which peters out around Brisbane (62). The Australian simulation also shows some connectivity between west-coast and east-coast warm temperate faunas via an eastward flowing extension of the Leeuwin Current and connectivity between eastern Australia and New Zealand (particularly the North Island) via the southwestern limb of the subtropical gyre, consistent with the distribution of *Aurelia* sp. 7 and biogeography (61). These results suggest a strong role for hydrography in shaping modern patterns of species diversity (5, 20, 63), which is a salutary finding in light of the growing emphasis, albeit debated and based largely on studies of fishes, on the role of behavior in generating or maintaining geographic structure in marine taxa (e.g., see refs. 6 and 64–66). Ultimately, extrinsic (e.g., currents) and intrinsic (e.g., behavior, life history) processes interact, with probably different relative effects on different larvae (6).

Consequently, as geographic isolation of marine taxa gains renewed attention (4–6), we need to develop tools that improve our ability to understand the interacting dispersive and isolating influences of ocean currents and animal behavior and their contributions to spatial patterns of genetic variation and evolution in marine taxa. Their influences may not always be obvious. Our global ocean model has demonstrated that even marine species with long planktonic stages, such as *Aurelia*, can be regionally restricted because of natural oceanographic patterns, thus having higher than expected geographic structure and species diversity. Therefore, relative to prior expectations, marine populations may be more susceptible to be threatened by, or to become, invasive species because of anthropogenic introductions that exceed the limited natural dispersal range.
We thank M. Baird, J. Carlton, B. Holland, L. Martin, and three anonymous reviewers for valuable comment. We also thank T. Bolton (Flinders University, Adelaide, Australia), J. Gardner (University of Wellington, Victoria, New Zealand), H. Ishii (Tokyo University of Marine Science and Technology, Tokyo), H. Miyake (Japan Marine Earth Science and Technology, Yokosuka, Japan), M. Omori (Akajima Marine Science Laboratory, Zamamison, Japan), P. Petersen (Mooloolaba Aquarium, Mooloolaba, Australia), S. Uye (Hiroshima University), H. Hiroshi-Hirama, and S. Wiclow (Tasmanian Agriculture and Fisheries Institute, Hobart, Australia) for contributing samples. This work was supported by a Vice-Chancellor’s Post-Doctoral Fellowship from University of New South Wales (to M.N.D.), an International Postgraduate Research Scholarship (to A.S.G.), and an Australian Research Council Large Grant (to M.H.E.).