**Population genetics of a recent range expansion in monarch butterflies**

Will Hemstrom1\*, Micah Freedman2,3\*, Myron Zalucki4, Sharon Strauss2,3, Santiago Ramírez2,3, Michael Miller1

1. Department of Animal Science, UC Davis
2. Department of Evolution and Ecology, UC Davis
3. UC Davis Center for Population Biology
4. School of Biological Sciences, University of Queensland, Australia

\*Authors contributed equally

**Introduction**

Understanding the patterns and processes that define species’ geographic ranges is a fundamental goal in ecology and evolutionary biology. Geographic ranges have historically reflected species’ fundamental niches, with changes in range size driven by gradual changes in climate (Hewitt 1999, Davis and Shaw 2001) and rare long-distance dispersal events (Gillespie et al. 2012). Over the past 200 years, the tempo and magnitude of species range expansions has greatly increased (Helmus et al. 2014, others). This is primarily the result of deliberate or accidental introductions associated with human agriculture and commerce, although many recent examples also highlight the role of anthropogenic climate change in mediating range shifts (Parmesan and Yohe 2003, others). Studying the population genetics of range-expanding species is important because it informs our understanding of important features such as the timing of expansion, the number of expansion / introduction events, and the amount of ongoing gene flow between ancestral and derived populations (Peter and Slatkin 2015).

Species that undergo range expansions often have distinctive patterns of population genetic structure, with decreasing relatedness and increasing contributions of genetic drift in populations further from the original source population (Hewitt 1996, Excoffier et al. 2009). One commonly encountered form of range expansion is stepwise serial stepwise dispersal, in which populations are founded in a stepping-stone fashion (Ibrahim et al. 1996, Slatkin and Excoffier 2012). Serial dispersal is characteristic of many post-glacial range expansions into temperate regions and has been shown for species as diverse as eider ducks (*Somateria mollissima*) (Tiedemann et al. 2004), ragwort (*Senecio halleri*) (Bettin et al. 2007), and rough-skinned newts (Taricha granulosa) (Kuchta and Tan 2005). The out-of-Africa expansion of *Homo sapiens* is also characterized by serial stepwise dispersal (Ramachandran et al. 2005, Liu *et al.* 2006).

One such species that has undergone a dramatic range expansion over its recent evolutionary history is the monarch butterfly (*Danaus plexippus*). Evidence suggests that the monarch historically occupied Central America and the southern United States before undergoing a large demographic expansion around 20,000 years (Zhan et al. 2014, Pfeiller et al. 2016). This demographic expansion likely coincides with the end of the last ice age and glacial retreat in North America, which enabled colonization of temperate areas by the monarch’s Asclepias host plants and likely set the stage for the onset of large-scale long-distance migration in the monarch. Much more recently, the monarch has become established around the globe in three independent out-of-North America expansion events (Zhan et al. 2014). This includes a southern expansion that involved establishment in South America and the Caribbean, an eastward expansion across the Atlantic and into the Iberian Peninsula, and a westward expansion across the Pacific. In this paper, we focus exclusively on the monarch’s range expansion across the Pacific.

Historical records suggest that the monarch crossed the Pacific quite recently, with the earliest positive records of monarch occurrences coming from the 1840s in Hawaii (Vane-Wright 1993, Zalucki and Clarke 2004). By 1871, the monarch had reached Australia and by 1900 was established on nearly every major Pacific island group. Some authors have attributed the recency of the monarch’s appearance in these locations to the “Columbus hypothesis,” which posits that the clearing of forests in the eastern and midwestern United States during the 1800s prompted a massive increase in the North American monarch’s population size and scope of migration (Vane-Wright 1993). However, demographic reconstructions using whole genome sequence data indicate that the monarch’s out-of-North America expansion events happened much more distantly, perhaps as long as 2,000-3,000 years ago (Zhan et al. 2014). Thus, there is still disagreement between demographic models and historical records about the timing of the monarch’s Pacific expansion.

Currently available population genetic data suggest that the monarch’s recent global range expansion happened in a serial stepwise fashion (Pierce et al. 2014a, Zhan et al. 2014). Serial expansion is indicative of natural expansion wave, rather than a series of independent and deliberate human introductions, as suggested by Zalucki and Clarke (2004). Still, there are a number of unsampled populations in the Pacific that might improve our understanding of establishment timing and direction.

Finally, in contrast to their migratory North American ancestors, nearly all Pacific island populations have become fully sedentary, year-round breeding populations. Little is known about how this contemporary loss of migration has affected fine-scale patterns of population differentiation in monarchs. One study has addressed this question: Pierce et al. (2014b) used microsatellites and showed that monarchs from Hawaii show little differentiation among islands. However, the conclusions of this study were based on only 11 variable loci from one island group. Thus, the degree to which loss of migration shapes fine-scale patterns of population differentiation remains unresolved.

In this study, we use reduced-representation whole genome sequencing across a sample of approximately 280 monarch butterflies to understand (1) patterns of relatedness among Pacific and North American populations, (2) expansion timing and amount of ongoing gene flow from North America, and (3) genetic differentiation within expansion populations. Our dataset contains tens of thousands of variable sites from monarchs in the ancestral North American population and many Pacific Island populations, including a number of previously unsampled locations: the Mariana Islands (Guam, Rota, and Saipan) and Norfolk Island. We find support for a stepwise pattern of dispersal across the Pacific, but with a previously uncharacterized westward expansion from Hawaii into the Mariana Islands. Estimates for the timing of the monarch’s establishment in the Pacific are concordant with a recent expansion, but with high uncertainty around the precise timing of this event. Gene flow from North America to Hawaii appears to be unidirectional, with low levels of ongoing North America to Hawaii gene flow. Monarchs within the Mariana Islands show strong patterns of differentiation despite being in extremely close proximity; by contrast, migratory North American samples form a single panmictic sample across the entire continent. Together, our data argue in favor of deferring to historical records to infer the history of the monarch’s range expansion and also provide a fascinating example of how migratory status can influence patterns of genetic isolation.

**Methods**

*Sample collection and storage*

Monarchs were collected as either larvae or adult butterflies from locations around the world between 1990 – 2017. When possible, monarchs were collected over as broad of a spatial and temporal window as possible to minimize the chances of sampling full or half sibs. Monarchs were stored in ethanol and keep either at room temperature or -20C freezers. For a summary of monarchs included in our sequencing, see Figure 1 and Table S1.

*Sample preparation*

*Data analysis*

**Results**

*Question 1: Overall patterns of relatedness*

PC1 explained (\_\_%) of the overall variance and separated North American from Pacific Island samples. PC2 explained (\_\_%) of variance and split Pacific Island populations into two out-of-Hawaii expansions. North American monarchs formed a single panmictic populations in all analyses.

STRUCTURE showed a generally similar pattern. The highest likelihood scenario was with k = 5. At k = 2, North American and Pacific Island populations were pulled apart. At k = 3, Guam was separately from all other Pacific population. At k = 4, samples from Rota were assigned their own cluster. At k = 5, Samoa, Fiji, and New Caledonia were assigned their own cluster, and at k = 6, Hawaii was assigned a cluster; k = 6 was the scenario with the highest likelihood. Values of k = 7 and higher only subdivided populations, and even at values as high as k = 10, eastern and western North American populations did not form discrete clusters.

*Question 2: Timing of establishment and patterns of ongoing gene flow*

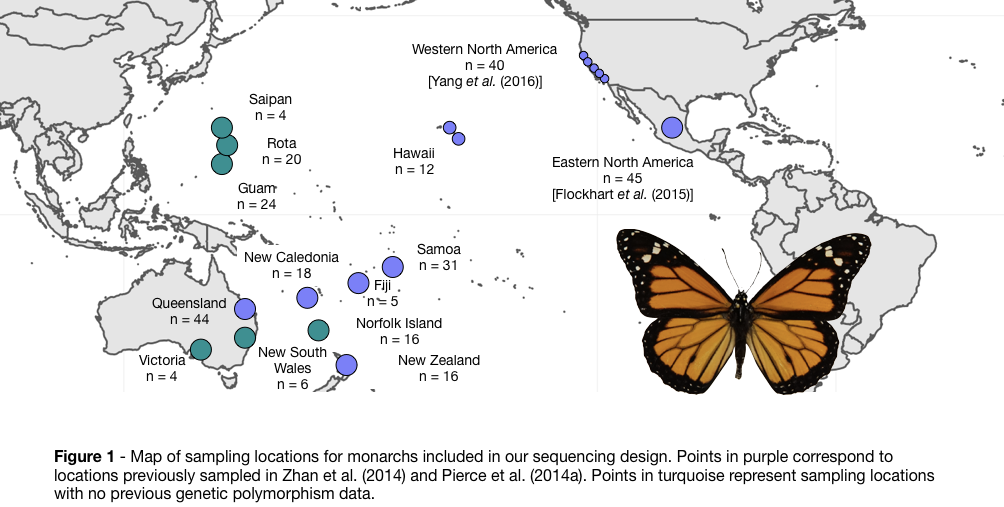
Models suggest a recent colonization of Hawaii from North America. However, the exact timing is difficult to pinpoint and is highly sensitive to model inputs, including the population growth model upon establishment. All models, regardless of introduction timing, agree on population expansion in both North America and Hawaii subsequent to establishment in Hawaii.

There appears to be a very small amount of ongoing gene flow from North America to Hawaii (--). However, there is no support for Hawaii > North America gene flow.

*Question 3: Patterns of differentiation within expansion populations*

Samples from the Mariana Islands (especially the well-sampled Guam and Rota populations) appear to form distinct populations, despite their extremely close physical proximity. By contrast, populations within Hawaii (Maui and Oahu) and Australia (Queensland, New South Wales, and Victoria) do not show strong patterns of differentiation. Likewise, we find support for a single panmictic North American population that encompasses both eastern and western populations.

**Discussion**



|  |  |  |  |
| --- | --- | --- | --- |
| **Population** | **Sampling Location** | **Sampling Year(s)** | **# Sequenced** |
| **North America** | Eastern North America (Mexican overwintering sites) | 2016 | 45 |
| Western North American (California overwintering sites) | 2015 | 40 |
| **Hawaii** | Maui | 2016 | 8 |
| Oahu | 2016 | 4 |
| **Mariana Islands** | Guam | 2015 | 24 |
| Rota | 2015 | 20 |
| Saipan | 2015 | 4 |
| **Fiji** | Viti Levu | 2009 | 5 |
| **Samoa** | Upolu | 2006, 2007, 2016 | 31 |
| **New Caledonia** | Grand Terre | 1991, 2006, 2010 | 18 |
| **Australia** | Queensland | 2016 | 44 |
| Victoria | 2016 | 4 |
| New South Wales | 2016 | 6 |
| **New Zealand** | North Island | 2007, 2011 | 6 |
| **Norfolk Island** | Norfolk Island | 2016 | 16 |
| **Total:** | | | **281** |
| Table S1 – Sampling locations and year for monarchs included in sequencing design | | | |