A tipping point in parapatric speciation

Ryo Yamaguchi*, Yoh Iwasa

Department of Biology, Faculty of Science, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan

ABSTRACT

More than two loci are involved in reproductive isolation in most cases of putative recent speciation. We study the speciation between two geographically isolated populations connected by infrequent migration, in which incompatibility is controlled by quantitative loci. Incompatibility genetic distance is defined as the fraction of compatibility controlling loci that are different between individuals. Speciation is established when genetic distance reaches a threshold level in spite of occasional migration and subsequent hybridization that reduce genetic distance. With stochastic analysis, we investigate how the time to speciation depends on the manner in which the magnitude of incompatibility increases with genetic distance. Results are: (1) The time to speciation is short if the migration rate is smaller than the mutation rate, or if intermediate levels of genetic distance cause mild incompatibility, making migrants less effective in reducing genetic distance. (2) Genetic distance may fluctuate around a positive quasi-equilibrium level for a long time, and suddenly show a quick passage to speciation when it goes beyond a "tipping point." Notably a gradual increase in incompatibility can result in a sudden and rapid formation of a new species. (3) Speciation becomes very slow if incompatibility is effective for individuals differing at only one locus. These findings provide testable predictions on reproductive traits controlled by specific incompatibility accumulation forms that facilitate the speciation process.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Speciation is the process in which a new species arises as reproductive isolation evolves between genetically diverging populations, and has been of great concern in evolutionary biology since Darwin (1859). Empirical studies suggest that the most frequent geographic mode of speciation is parapatric or closely allopatric, because different populations are indeed isolated, but not absolutely with infrequent migrations (Endler, 1977; Barton and Hewitt, 1981). Therefore, hybridization occurs almost inevitably during speciation (Gavrilets, 2004; Fitzpatrick et al., 2008; Abbott et al., 2013).

In the Biological Species Concept (Dobzhansky, 1935; Mayr, 1942), reproductive isolation (or barriers to gene flow) was found to play a key role in the definition of species. Prezygotic isolation is caused by differences in mating behavior or gametic recognition that may reduce the success rate of mating or fertilization. Postzygotic isolation due to sterility and lethality of hybrids is also widely observed (Mayr, 1963; Coyne, 1992). The genetic basis of low hybrid fitness typically involves interactions between different loci, and low fitness may be caused by alleles that yield high fitness in their usual genetic backgrounds (Orr, 1995; Johnson and Porter,

2000; Turelli and Orr, 2000). From a literature review, Coyne and Orr (1998 , 2004) concluded that the number of loci controlling the incompatibility varies substantially between species and between the focal traits (viability, sterility, male courtship, sexual traits, ecological traits, pheromone, etc.). In many cases of putative recent speciation, often more than two loci, and sometimes more than 10, are involved in reproductive isolation. Incompatibility between two diverging populations not only appears to be driven by ecological adaptation (Schluter, 2009) but may also result from responses to purely mutational mechanisms or to internal genetic conflicts (Orr and Presgraves, 2000; Maheshwari and Barbash, 2011). Both prezygotic isolation and postzygotic isolation reduce effective gene flow, thus contributing to speciation.

Theoretical studies have been performed on the process of establishment of reproductive isolation (speciation) under a given genetic system controlling incompatibility. The Dobzhansky–Muller (DM) model (Dobzhansky, 1937; Muller, 1939 , 1942) is one such model used for this purpose. In this model, individuals differing only at one locus can cross freely, allowing neutral accumulation of separate mutations in different populations. Eventually, two populations become reproductively isolated because individuals differing in both loci cannot produce viable offspring. This classical model highlights the role of epistasis in speciation. Various extensions of this model have been examined in relation to the process of allopatric speciation (Higgs and Derrida, 1992;

* Corresponding author.
E-mail address: ryamaguchi@bio-math10.biology.kyushu-u.ac.jp (R. Yamaguchi).

Manzo and Peliti, 1994; Gavrilets and Gravner, 1997). In parapatry, Navarro and Barton (2003) found that chromosomal rearrangements could promote the accumulation of such incompatibilities. To overcome gene flow, Agrawal et al. (2011) explored the origin of DM incompatibility where genes subject to natural selection also affect intrinsic isolation, either directly or via linkage disequilibrium. Bank et al. (2012) also investigated that how much gene flow was needed to inhibit the accumulation of DM incompatibility. These researches emphasize that mechanisms to overcome gene flow are important for species formation.

Extending DM model to multiple loci is straightforward to consider graduated incompatibilities. We may consider the situation in which a number of loci scattered throughout the genome control incompatibility (Wu and Palopoli, 1994; Naveira and Maside, 1998). Two individuals differing in some fraction of the loci can cross without problems if the genetic difference is less than a certain threshold, but cannot produce viable offspring if the difference is equal to or greater than that threshold (Turelli and Orr, 1995; Gavrilets, 2000; Yamaguchi and Iwasa, 2013, 2016 , 2017). This becomes the DM model when the number of incompatibility-controlling loci is equal to two. An additive incompatibility model is an even simpler quantitative genetic model (Gavrilets, 2000). The magnitude of the incompatibility caused by each locus is small but different loci combined in an additive or multiplicative manner can result in incompatibility between individuals (Walsh, 1982; Orr, 1995; Orr and Turelli, 2001; Gavrilets, 2004; Welch, 2004; Livingstone et al., 2012).

The magnitude of incompatibility between two populations increases with genetic differentiation, but if they are perfectly isolated (no migration), then the time to speciation (i.e., the time for the genetic difference to reach a certain threshold level) depends only on whatever processes drive substitutions within populations. However, rare but recurrent migration occurs between two populations, as a byproduct of storms or other accidental occurrences. In the presence of recurrent migration, the genetic difference between two populations does not increase monotonically with time; its dynamics is determined by the balance between the increase due to accumulation of separate mutations in different populations and the decrease due to migration and subsequent hybridization (Gavrilets, 2000). Hence, the magnitude of incompatibility for an intermediate level of genetic differentiation is very important in determining speciation dynamics.

The mechanism by which the incompatibility increases with the number of different loci is not known. However, there are a number of reports on how the magnitude of incompatibility changes with the divergence time between two closely related species or geographically isolated populations of the same species (subspecies). Incompatibility increases with time at a constant rate (linear effect) or at a decelerating rate (slowdown effect, cf. Orr, 1995) in fungi (Giraud and Gourbière, 2012) and in cichlid (Stelkens et al., 2010). Gourbière and Mallet (2009) reviewed several combinations of substitution processes and incompatibility interactions in various taxa, and found that cases in which incompatibility increases with time at an accelerating rate (snowball effect) are rare. Palmer and Feldman (2009) tried to explain these by modeling the genetic network underlying the incompatibility trait. The classical DM incompatibility is also recognized as rare in Drosophila (Brideau et al., 2006). However, some recent studies offer empirical evidence for a snowball effect of postzygotic incompatibilities in Lepidoptera hybrid viability (Presgraves, 2002), centrarchid fishes (Bolnick and Near, 2005), Drosophila (Matute et al., 2010), and Solanum (Moyle and Nakazato, 2010; Städler et al., 2012). In general, incompatibilities may be expressed by complex combinations of cumulative genetic changes and their gene regulatory networks. Many genes contribute to incompatibility, and nonadditive genetic effects on hybridization efficiency have

been observed in Solanum (Moyle and Nakazato, 2008), Arabidopsis (Burkart-Waco et al., 2012) and Drosophila (Corbett-Detig et al., 2013).

In this paper, we focus on the time course to parapatric speciation between two populations of a species with recurrent migrations when the magnitude of incompatibility between two individuals gradually increases with genetic divergence. Our model allows one to treat both ecological incompatibility and intrinsic incompatibility within the same framework. These populations must overcome gene flow by migration to accomplish complete speciation. Under the assumption that populations are close to monomorphic most of the time, we examine how the time to speciation depends on the shape of the relation between genetic differentiation and incompatibility. The analysis shows that the time to parapatric speciation strongly depends on the genetic distance-incompatibility relations and that reproductive traits controlled by specific incompatibility accumulation forms can facilitate speciation.

2. The model

For simplicity of analysis, we here consider a haploid species that engages in sexual reproduction with discrete and non-overlapping generations. Consider two finite populations, each comprising N individuals. The populations live on two islands, or on island-like habitats separated geographically. Very rarely, for example in storms, a group of adult individuals from one population may migrate to and land on the second population and participate in reproduction processes there. In this section we introduce the procedures of individual-based simulation.

2.1. Incompatibility genetic distance

Incompatibility is controlled by l loci on autosomal chromosomes. The sequence of autosomal l loci of individual α is represented as $L^\alpha = (L_1^\alpha, ..., L_l^\alpha)$, where L_i^α is the allelic state of the ith locus. Initially, all the individuals have the same state for all loci. Mutation occurs at each locus with the mutation rate u per generation. We assumed an infinite allele model (Kimura and Crow, 1964), implying that novel mutations are different from all the existing alleles. We also assume free recombination. Let us define the "incompatibility genetic distance" (or simply, the "genetic distance") z between individuals as the fraction of incompatibility-controlling loci that differ between them, which is equal to the normalized Hamming distance. This is analogous to the number of segregating sites in a sample of two gametes, which is widely used in molecular evolutionary genetics (Li, 1997; Cruickshank and Hahn, 2014), and to the number of heterozygous loci in a diploid organism (Gavrilets, 1999). For genetic distance between polymorphic populations, we use average genetic distance calculated from all combinations of individuals in each population.

We adopt the Wright–Fisher Model for random genetic drift (Fisher, 1930; Wright, 1931; Crow and Kimura, 1970). In this model, parents are randomly chosen from the population (mother and father cannot be the same individual) with replacement, and a descendant is made from its parents' loci chosen with equal probability.

Migration occurs very infrequently following a Poisson process at a rate of m per generation. There is no simultaneous bidirectional migration because the rate m is very small. At each migration event, constant N' adults may migrate as a bunch before mating and participate in reproduction on the other island. N' is much smaller than the resident population size N. These immigrants are chosen from the source population randomly, and they participate in mating. The selected immigrants are removed from their own population, and the remaining individuals produce the next gener-