This is an electronic appendix to the paper by Kawata 2002 Invasion of vacant niches and subsequent sympatric speciation. *Proc. R. Soc. Lond.* B **269**, 55-63.

Appendix A

Simulation results for individuals using more than one food type

Individuals use more than one food type, especially if the types are quantitatively different, and with small differences between close types. Therefore, the case in which an individual uses more than one food type was also considered.

The standard deviation σ for resource-use was varied, from 0.4 to 0.8 (Fig. A1). For instance, setting σ=0.6, if an individual's resource-use phenotype is 8, the individual uses food types 6, 7, 8, 9 and 10 with probabilities of 0.01, 0.20 0.58, 0.20, and 0.01, respectively. Suppose that the individual searches for foods within the 25 cells of the area defined by x=60-100, y=170-220. Within this area, the cells containing food types 6, 7, 8, 9 and 10 number 0, 2, 4, 5 and 0, respectively. Thus the foraging success of the individual is F=0 x 0.01 + 2 x 0.2 + 4 x 0.58 + 5 x 0.2 + 0 x 0.01 =3.72.

When individuals used more than one food type, as shown in Fig. A1, individuals with different resource-use phenotypes could consume the same food types. Such competition through consuming food types was also taken into account for calculating *N*. Thus,

$$N = \sum_{i=1}^{n} I_i$$

where I_i is a relative weight for individual *i*, and *n* is the number of individuals whose resource-use overlaps those of the focal individual. I_i for an individual *i* is the sum

$$\sum_{u=-f}^{f} D_u E_u$$

where D_u and E_u are the probabilities of use of food type u by the focal individual and the competitive individual, respectively, and f is the range of resource-use for the focal individual. For instance, when the standard deviation for resource-use (□) is 0.4 (Fig. A2), and $f = 1$, an individual with resource-use phenotype 8 uses food types 7, 8 and 9 with probabilities 0.1, 0.8 and 0.1, respectively. Then the I value of an individual with resource-use phenotype 7 as a competitor of the focal individual with resource-use phenotype 8 is 0.16 (=0.1 x 0.8 + 0.8 x 0.1 + 0.1 x 0). For the focal individual, probabilities for using food types 7 and 8 are 0.1 and 0.8, respectively, whereas for the individual with resource-use phenotype 7, they are 0.8 and 0.1, respectively. Therefore, one individual with resource-use phenotype 7 reduces the focal individual's amount of available food by the relative amount 0.16.

Fig. A2 shows the probabilities for the four outcome categories when individuals could use more than one food type. When genetic neighborhood size was small, the probability of complete speciation (I) was smaller when the standard deviation □of resource-use for individuals was 0.4 (i.e., an individual used more than one food type) than when □ was 0 (i.e., an individual used only one food type). However, thereafter the probability of complete speciation increased with increasing values of □.. More specifically, for □=0.6 and □=0.8, small genetic neighborhood size augmented complete (I) or incomplete speciation (II), and resource-use evolved by diverging to the values of the two sets of most abundant food types (Fig. 9, C). In contrast, when genetic neighborhood size was large, complete speciation was not observed regardless of the value for □ (i.e., □=0, 0.4, 0.6, 0.8).

Two factors might pertain to the effect on speciation of resource-use variation for individuals (Fig. A2). First, when an individual can use more than one food type, individuals who use mainly the intermediate less abundant food type can also use other more abundant food types. Disruptive selection may thus become weak, and also then, the probability of speciation. Second, when the range for use of food types is larger, competition among individuals with different resource-use phenotypes became more intense. Through such exploitative competition, individuals using most abundant food types will be more likely to out compete other individuals (Fig. A2, C). This facilitates disruptive selection on resource-use phenotypes, which leads to increase in the probability of complete speciation (Fig. A2, A).

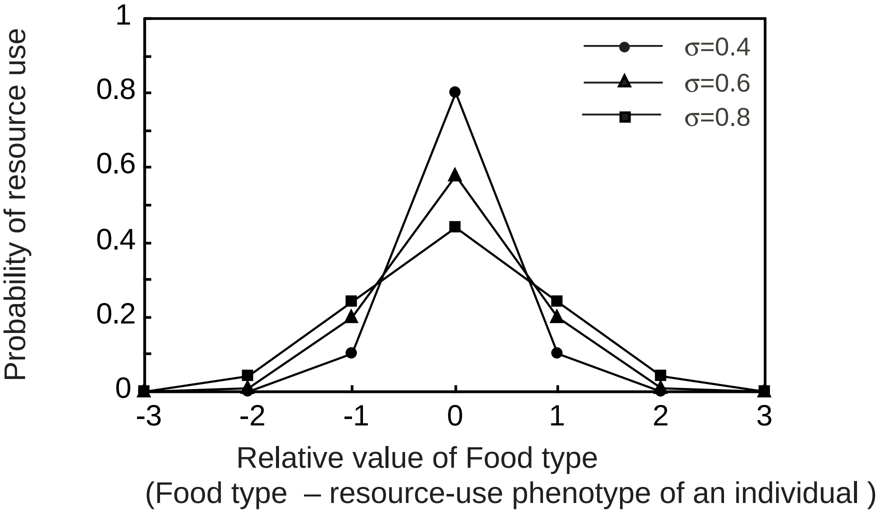

FIG. A1. The probability of use of different food-types by an individual (when individuals use more than one type). The individual's preferred food-type, corresponding to its resource-use phenotype value, is given the value 0 (in relative terms) on the horizontal axis; and other adjacent food-types, with higher (minus) and lower (negative) values, fall to the left and right, respectively, of the central value.

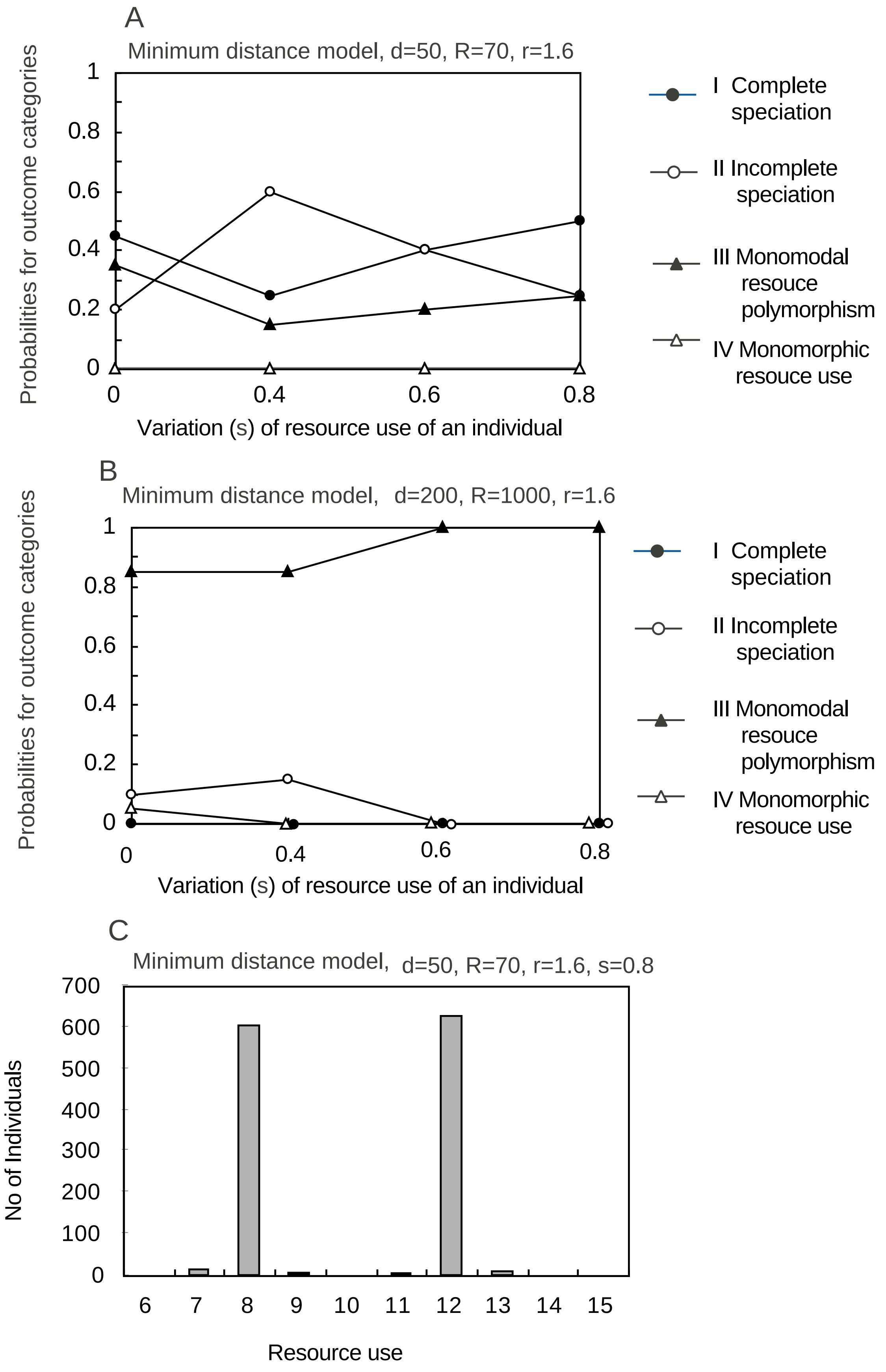

Fig. A2. The effects of resource-use variation of an individual on probabilities for realizing the four outcome categories. The probability of resource-use by an individual for different values of was shown in Fig.A1. In terms of numbers of individuals using the different food types, Part C of the figure shows the outcome of evolution of resource-use for one simulation when complete speciation was observed (□=0.8).

Appendix B

Measurements of Spatial Structure

For describing the spatial distribution of individuals, normally the number of individuals is calculated for quadrates of defined size to obtain a dispersion index. Here, however, Morishita's index of dispersion (Morishita 1962; Krebs 1999) was used. The number of individuals within a circle is counted whose center is the location of a focal individual. The radius of the circle was set at 50, 70, or 90. Morishita's index of dispersion ($I_{□}$) is calculated as

$$I_{\delta} = n\left[\frac{\sum_{i=1}^{n} m_i^2 - \sum_{i=1}^{n} m_i}{\left(\sum_{i=1}^{n} m_i\right)^2 - \sum_{i-1}^{n} m_i}\right]$$

where m_i is the number of individuals within the circle of focal individual i, and n is the number of the focal individuals. The spatial distribution of individuals is clumped, random, or uniform if $I_{□} > 1$, $I_{□} = 1$, or $I_{□} < 0$, respectively.

The spatial distribution of phenotypes was examined using semivariance (Robertson 1987; Burrough 1995). When the measured values of a variable vary at different spatial points, the variable can be written as $z(x_i)$, where z represents the spatial variable and x_i the spatial location i. The semivariance (□ (h)) for a spatial distance h was estimated by

$$\gamma(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}\left[z(x_i)-z(x_{i+h})\right]^2$$

where $N(h)$ is the total number of sample point contrasts for the distance in question. Here, the phenotypic values of resource-use are spatial variables. Spatial distance h is divided into 25 interval-classes, each interval being 20 units (of the W x W square). Thus, h ranged from 0 to 500. The expression $z(x_i) - z(x_{i+h})$ means the difference in phenotypic value between an individual located at *i* and an individual whose location is a distance h±10 from *i*. When there was more than one individual located a distance h±10 from *i*, differences in phenotypic values between all the individuals located at that distance from *i* were calculated.

The graph plotting semivariance (vertical axis) versus spatial distance h (horizontal axis) is called a semivariogram. The semivariance rises gradually and then levels out. The values of the horizontal axis (*h*) and the vertical axis (□ *(h)*) at which the semivariance levels out are called the range and the sill of the semivariogram, respectively. Because the range is the interval of values of h at which the spatial distribution of individuals with similar or same values of resource-use phenotypes tend to be clumped and contiguously distributed, the values of the range are equivalent to the length of the clumps of phenotypes with similar values, and the sill represents the spatial variability of the values of resource-use phenotype.

Part A of Fig. A3 shows the dispersion index for one simulation that resulted in speciation (at about the 1100th generation). The dispersion index was slightly higher than 1, but not significantly different from 1 ($P > 0.05$). Thus, at least in this case, spatial distribution was nearly random. The average number of males within a circle of radius 50, 70, and 90 was 7.51, 14.89, and 23.92, respectively. For the same simulation, Part B of Fig. 10 gives the semivariograms. This showed that with increasing distance (*h*) to about 100, the semivariance increased, which indicated that the same or similar phenotypic values (resource-use) tended to clump into contiguously distributed sets within this spatial scale. Part C of Fig. A3 shows one example of the spatial distribution of individuals using smaller food sizes and larger food sizes at the 500th generation, which was before the occurrence of complete speciation at the 1100th generation.

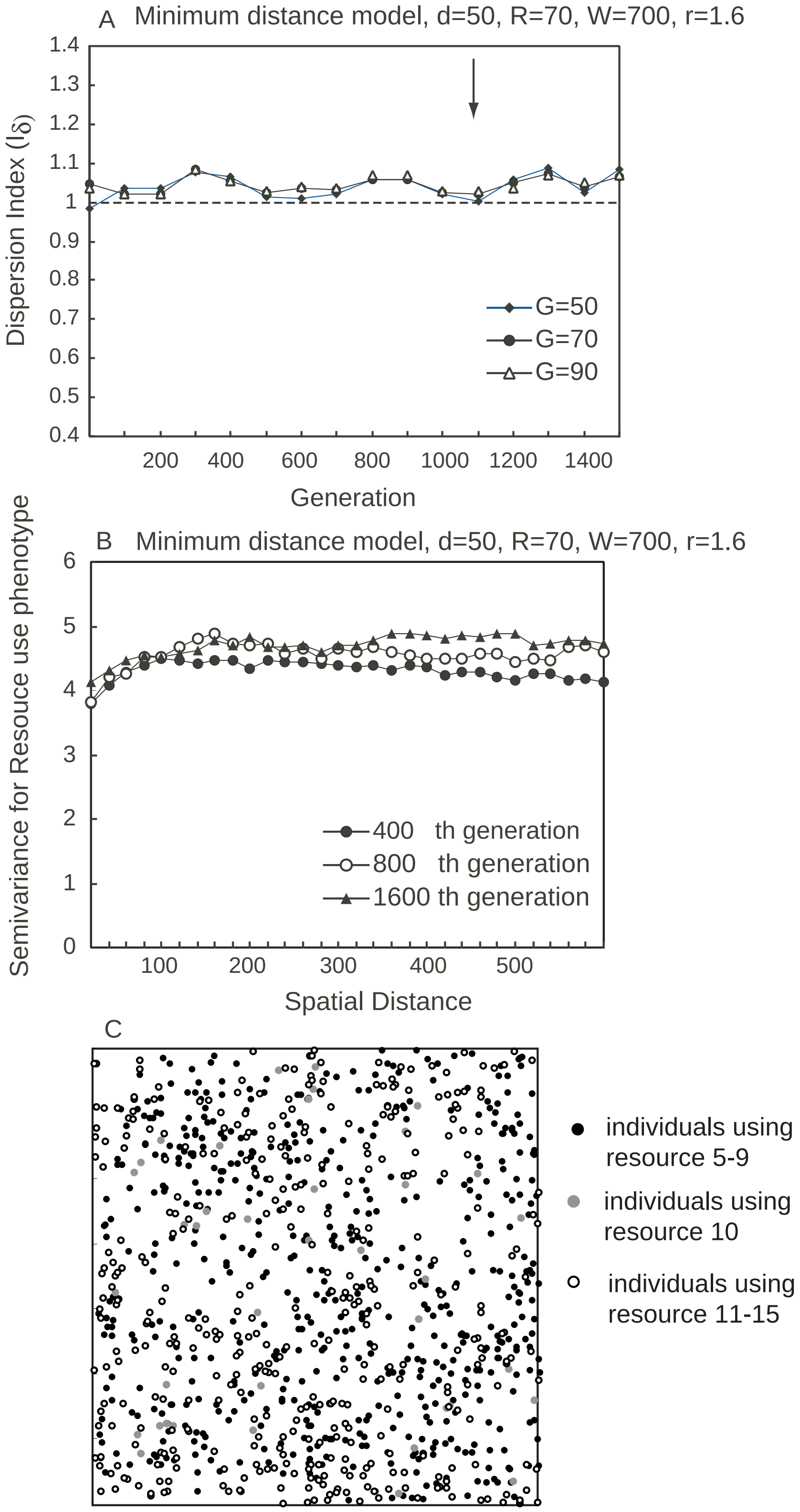

Fig. A3. Part A: Dispersion index. Here G is the radius of the circle within which the number of individuals was counted for the dispersion index. The arrow indicates the time when reproductive isolation was completed. For the simulation used in this graph, speciation was completed at the 1100th generation. Part B: Semivariogram. Part C: Spatial distribution for resource use for one simulation, which resulted in complete speciation. For all parts of the figure, the initial number of loci for resource-use, male sexual trait, and female preference was 4. The resource distribution was R1.

References

Burrough, P. A. 1995 Spatial aspects of ecological data. In *Data Analysis in community and landscape ecology* (eds. R. H. G. Jongman, C. J. F. Teer Braak & O. F. R. Tongeren), pp. 213-251. Cambridge: Cambridge University Press.

Krebs, C. J. 1999. *Ecological methodology*. Menlo Park: Addison-Welsey.

Morishita, M. 1962 Id-index, a measure of dispersion of individuals. *Res. Popul. Ecol.* **4**, 1-7.

Robertson, C. P. 1987 Geostatistics in ecology: interpolating with known variance. *Ecology* **68**, 744-748.

Appendix C

The effects of size of habitat area (W), reproductive rate on the probability of speciation

When reproductive rates (r) were larger than 1.5, the probability of complete speciation (I) became larger (Fig. A4). Similarly, when the habitat area (W x W) was increased from 500 x 500 (area =250000) to 700 x 700 (area=4900) or 900 x 900 (area =810000), the probability of complete speciation increased (Fig. A4, B).

The probability of speciation was also affected by reproductive rates and the size of the habitat area (Fig. A4). If the reproductive rates were small, the probability of a female finding a mate with an acceptable or optimal phenotype, similar to her own, is somewhat decreased. Also, chances for mutants to arise that can invade the vacant niche are reduced. For both of these reasons, lower reproductive rates might cause less frequent outcomes of complete speciation.

Often, when the size of the habitat was small (W=500), few or no individuals using the larger sizes of food were present in the initial population, because the number of total individuals was low. Thus, when W was 500, the possibility of invading the vacant niche of larger sizes of food was less.

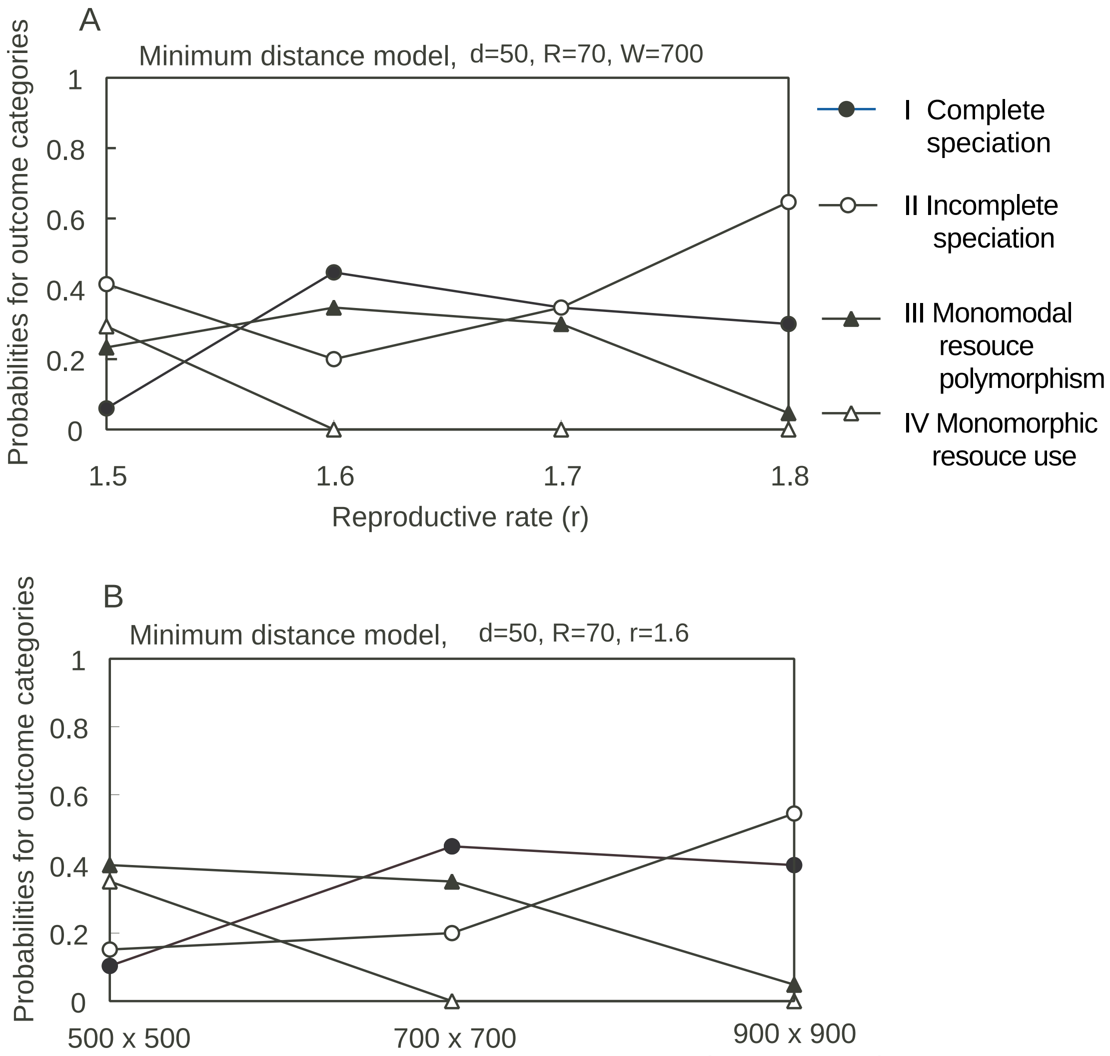

FIG. A4. The effects of reproductive rate and habitat size on probabilities for realizing the four outcome categories. The initial number of loci for resource-use, male sexual trait, and female preference was four. Resource distribution was R1.