

## Supplementary Information

We present here a more comprehensive discussion of the network model, its assumptions, its limitations and the sensitivity of our results to changes in the parameters. Of particular interest is the effect of long-range flea movements on the model behaviour. We also provide results on model sensitivity to changes in (i) the weighting given to the flea's own burrow system when determining the destination of a dispersing infectious flea ( $C$ ), and (ii) the rate at which a family group recovers from plague relative to the rate at which infectious fleas disperse from infected burrow systems.

The model assumes that all family groups of great gerbils are of equal size and hence equally susceptible and equally infectious if infected. This is a simplification since the size of a family group varies seasonally and may be correlated with the overall density of great gerbils such that at high occupancy family groups tend to be larger. If larger family groups are both more susceptible and more infectious such that these sources of heterogeneity are correlated then this may affect the model behaviour, however one can show that if they are not correlated then such heterogeneity will not lead to qualitative changes in the spread<sup>1-3</sup>. We did not attempt to model this as while such correlations are biologically plausible there is no field data to support them or on which to base more complex modelling.

The burrow system density of great gerbils across central Asia is highly variable, within plague foci and between foci. Burrow system density was measured as part of the plague surveillance initiated in Kazakhstan in the 1940's and estimates from the PreBalkhash focus vary from 0.89 to 4.8 burrow systems per hectare. The substantial spatial variation is thought to be caused by differences in the plant community and the suitability of the soil for digging. The effect of such variation is predictable from our model though, given that the distribution of weights used to determine flea movements between burrow systems is fixed. This is because model behaviour is actually determined by the product of burrow system density and the fraction occupied, i.e. by the density of *occupied* burrow systems, since empty burrow systems play no role. This means that varying the density of burrow systems is equivalent to rescaling the horizontal axis in Fig. 2 of the main text (representing the abundance of great gerbils as the fraction of burrow systems occupied). For example, an occupancy rate of 0.25 when the burrow system density is 3 per hectare is equivalent to an occupancy rate of 0.5 when the burrow system density is only 1.5 per hectare. This implies burrow system density will change the value of the threshold and since the

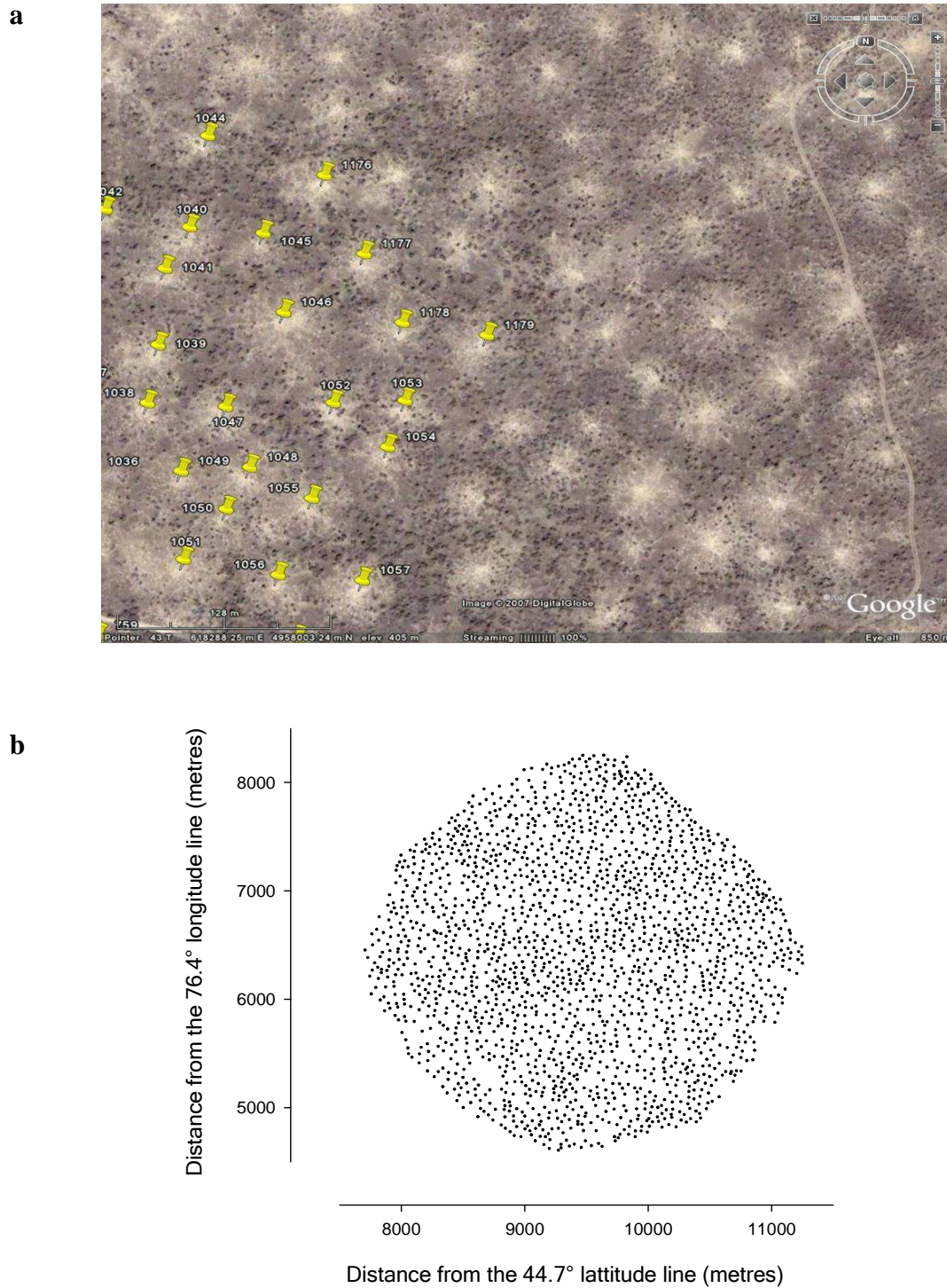
rescaling can either stretch or compress the axis, it may also affect the abruptness of the threshold. For example, we would predict that a decrease in burrow system density will tend to both raise and soften the threshold. Spatial heterogeneity in burrow system density at the sector scale (10km × 10km) is also likely to exist but has not yet been studied. Intuitively, areas of low burrow system density are barriers to the spread of plague while high burrow system density assists spread of plague. An interesting feature of the landscape in the PreBalkhash focus is the presence of ancient river beds that appear to have high densities of burrow systems. There are now ongoing field studies to collect ground truthing data for image analysis that will produce maps of burrow system density. One application of such maps will be to verify more carefully the extent to which the ancient river beds have a burrow system density that is distinctly different from the density in the surrounding landscape.

We now turn our attention to the distribution of flea movements used to determine which burrow system an infectious flea arrives at when a transmission event occurs in the network model. The mark-recapture data used to construct the distribution of great gerbil movements, on which the distribution of flea movements is based, has been described in detail elsewhere<sup>4,5</sup>. We emphasise here that we do not assume that the distribution of flea movements is independent of burrow system density. When determining where infectious fleas will disperse to, then the weights given to each occupied burrow system in the neighbourhood (see Methods) have the same dependence on distance regardless of occupancy levels. This is not the same as assuming a fixed distribution of gerbil/flea movements and the effect is that at low occupancy the flea movements tend to be larger than at high occupancy. To judge the sensitivity of the model we considered two additional distributions; (i) a geometric distribution (having a thicker tail), and (ii) another half-logistic distribution obtained by assuming the mark-recapture data sampled only 95% of the relevant movements transporting fleas and that the remaining 5% had a mean of 1km. The three distributions are shown in Fig. S2 and the results of simulations are shown in Fig. S3. For both distributions the radius determining the neighbourhood of infected burrow systems was also extended to 1.5km. The qualitative behaviour was unchanged and  $R_0$  remained 1.5 at the percolation threshold, though the location of the threshold for the second alternative distribution decreased to ~0.17. This distribution represents an extreme as it implies that 1 in 20 flea movements are around 1km whereas the details of two flea studies<sup>6,7</sup> suggest it is at most 1 in 100 that are this distance. In one study 3 movements of ~1km were observed, 1200 m was the maximum and 2000 fleas were recaptured, while in the second 500 fleas were recovered and 2 movements of length ~1km were observed. The search radius for both studies was around 2km.

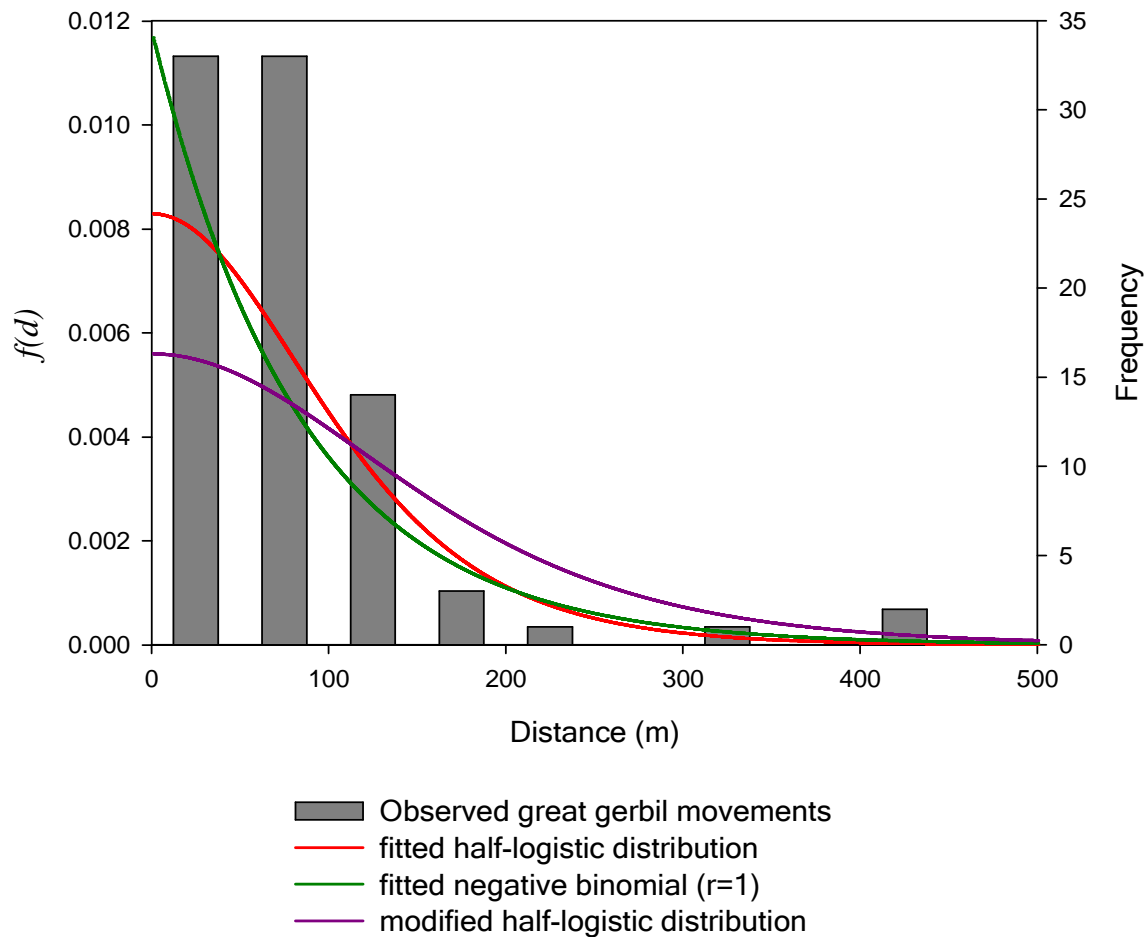
We note again here that we measured  $R_0$  directly from the network simulations by recording the number of vertices infected by the first infected vertex and then averaging over many simulations. In fact, it is difficult to define  $R_0$  in any other way, because, whereas in well-mixed and very large host populations, as an epidemic proceeds, the size of the  $k$ th generation of infecteds relative to the size of the  $(k-1)$ th generation tends to  $R_0$ , for spatial epidemics, this ratio tends to 1. We hasten to add that a spatial epidemic is still like any other epidemic, though, in that the expected number of secondary infections arising from the very first infection must be greater than 1 in order for the epidemic to start. Hence,  $R_0 > 1$  remains a necessary condition for epidemic spread though it is not a sufficient condition.

While the networks we studied were random, we note that the spatial arrangement of the family groups may not be random and may be very different from the regular spatial arrangement of the burrow systems seen in Fig. 1 of the main text. The available field data suggests it is random but such data are only available for small areas (the two field sites where mark-recapture studies were done were  $500 \times 600$  m and  $500 \times 500$  m) and there are several biologically plausible alternatives. The most interesting of these is that great gerbils may have a tendency to colonise empty burrow systems that are close to other family groups creating clusters of occupied burrow systems.

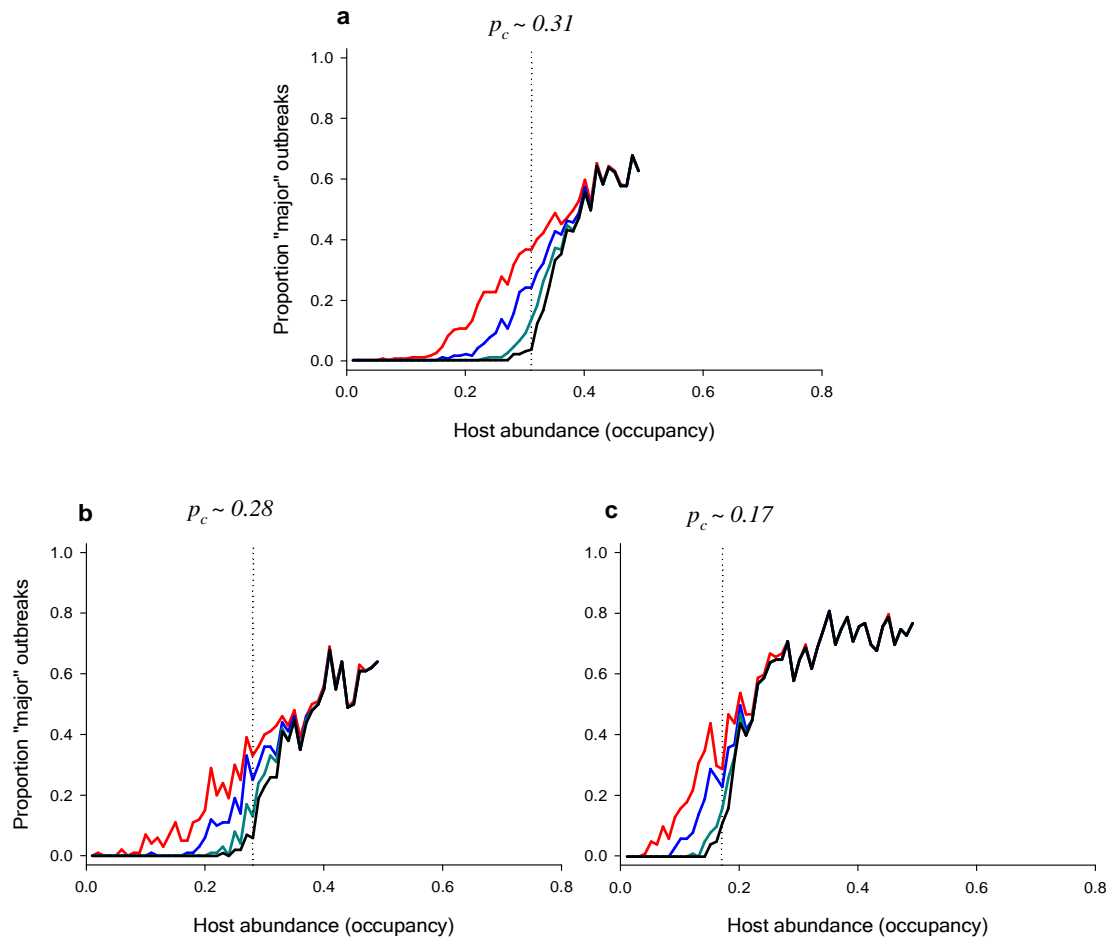
In Fig. S4 we show the results of the network model when the relative recovery rate and the parameter  $C$  are varied by  $\pm 10\%$ . The results suggest that the value of the percolation threshold is more sensitive to the rate of recovery (relative to the rate of transmission events) than it is to changes in  $C$ . The results also suggest that the sensitivity to the relative recovery rate is asymmetric in the sense that an increase in the recovery rate (a ratio of 1:9) caused a larger change in the percolation threshold than a decrease (a ratio of 1:11). In all cases the direction of the shifts in the percolation threshold were predictable in the sense that faster recovery rates predictably increased the threshold while slower recovery rates decrease the threshold.



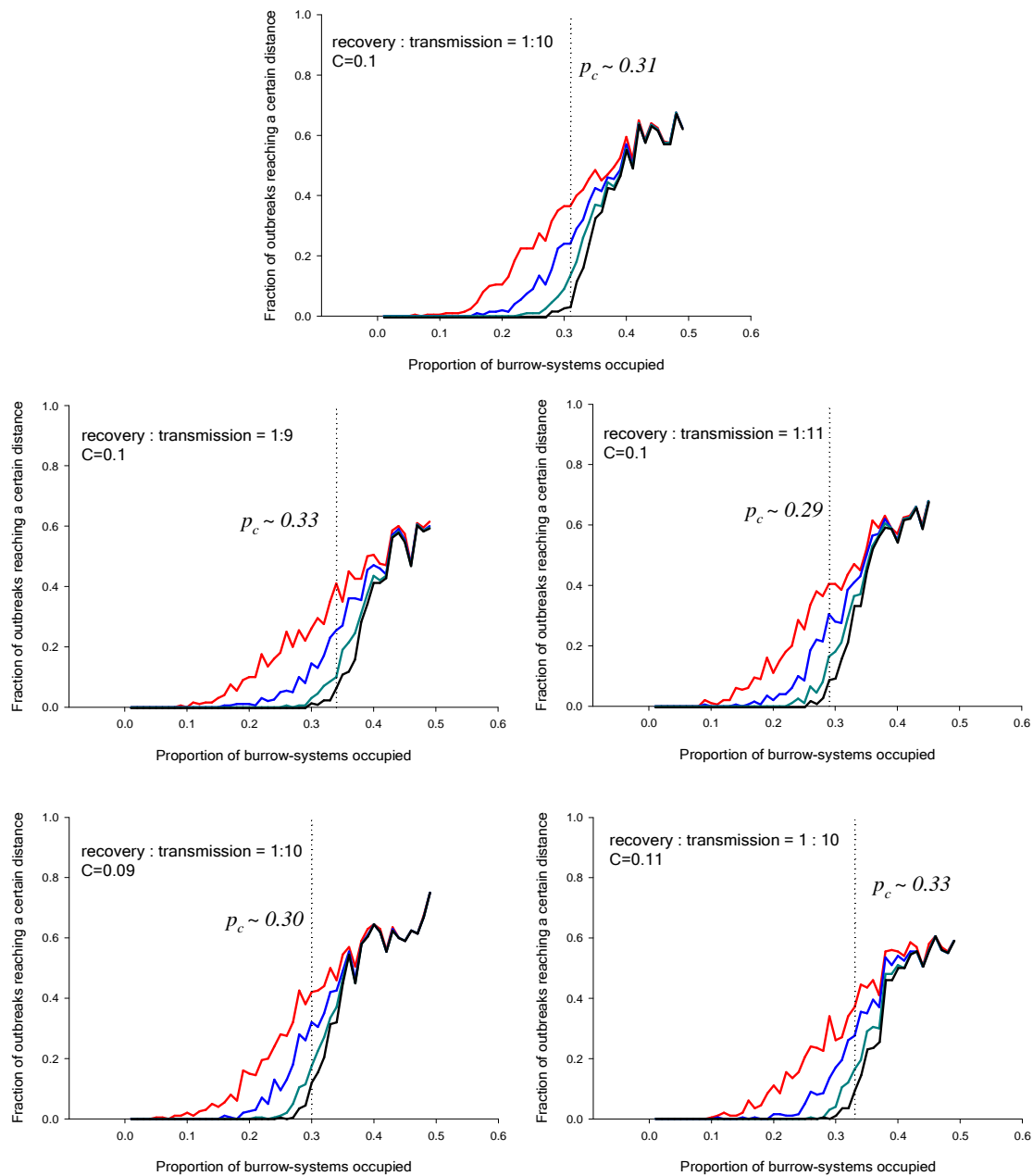
**Figure S1.** (a) An image captured using the software Google Earth (<http://earth.google.com/>), showing the marking of burrow systems and recording of coordinates. (b) The resulting set of points plotted as distances (in metres) from the 76.4° longitude line and the 44.7° latitude line.



**Figure S2.** A histogram of the movement distances ( $n=87$ ) of great gerbils recorded during two years of mark-recapture field studies, shown together with several probability distribution curves used for the function  $f(d)$  to construct weights that govern the destinations of infected fleas. The three distributions are the half-logistic distribution fitted with the mean calculated from the observed great gerbil movements (83.57 m), the geometric distribution (with the same mean) and a second half-logistic distribution obtained by adding four movements of length 1km and recalculating the mean (which shifted to 123.8 m). Transmission of plague between family groups relies on infectious fleas dispersing to nearby burrow-systems which they achieve by migrating to burrow system entrances and jumping to a passing animal, usually a great gerbil. With this distribution 93% of gerbil movements between burrow systems, and hence flea movements between family groups of great gerbils, are less than 200m.



**Figure S3.** Network model results for the three distributions shown in figure S2; the original half-logistic distribution with mean equal to the mean of the recorded great gerbil movements (83.57 metres) (a), the geometric distribution (also with mean set to the observed mean of great gerbil movements) (b) and the half-logistic distribution obtained by adding four movements of length 1km and recalculating the mean (123.8 metres) (c). The curves represent the fraction of simulations resulting in new infections 750 (red), 1500 (blue), 3000 (cyan) and 4500 (black) metres from the site of initial infection. The number of simulations for each value of occupancy for (a) was 200, but only 100 for (b) and (c) and for these two the neighbourhood of an infected burrow system was also extended to 1500 metres.



**Figure S4.** Network model results for various values of the parameters  $C$  (the weighting given to the flea's own burrow system when determining the destination of a dispersing infectious flea) and the ratio of recovery events to flea dispersal (transmission) events. Curves represent the fraction of simulations resulting in new infections 750 (red), 1500 (blue), 3000 (cyan) and 4500 (black) metres from the site of initial infection.

## Reference List

1. Trapman,P. On stochastic models for the spread of infections. 2006. Vrije Universiteit Amsterdam.  
Ref Type: Thesis/Dissertation
2. Mollison,D. The dependence of epidemic and population velocities on basic parameters. *Math. Biosci.* **107**, 255-287 (1991).
3. Kuulasmaa,K. The spatial general epidemic and locally dependent random graphs. *Journal of Applied Probability* **19**, 745-758 (2007).
4. Begon,M. *et al.* Epizootiological parameters for plague (*Yersinia pestis* Infection) in a natural reservoir in Kazakhstan. *Emerg. Infect. Dis.* **12**, 268-273 (2006).
5. Davis,S. *et al.* Plague metapopulation dynamics in a natural reservoir: the burrow-system as the unit of study. *Epidemiol. Infect.* **135**, 740-748 (2007).
6. Korneyev,G.A. The quantitative characteristics of parasitic exchange in some mammals species resulting from simulation of epizootics in desert biocenoses. 1968. Microbe Institute, Saratov.  
Ref Type: Thesis/Dissertation
7. Rudenchik,Yu.V., Soldatkin,I.S., Severova,E.A., Mokiyeovich,N.A., & Klimova,Z.I. Quantitative evaluation of the possibility of the territorial advance of epizooty of plague in the population of the *Rhombymus opimus* (northern Karakum). *Zoological Zhurnal* **46**, 117-123 (1967).