Introduction

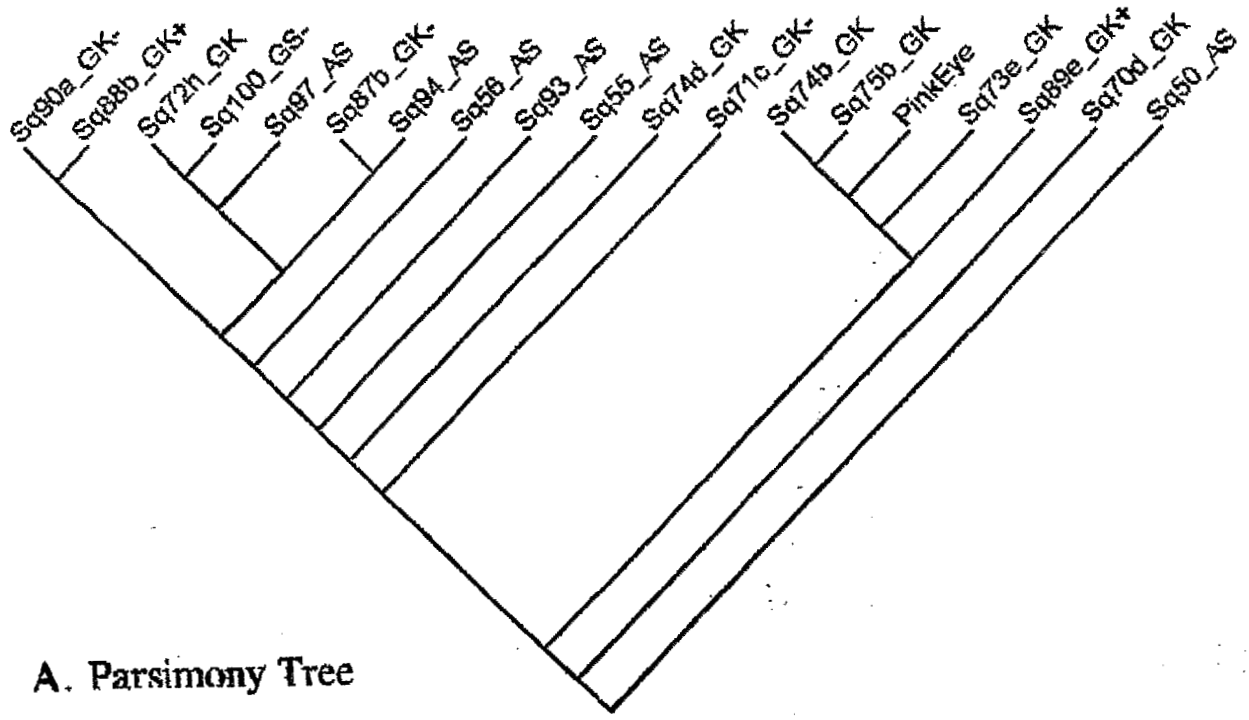
We are interested in the phenomenon of polymorphic chromosomal inversions as a potential force structuring variation in natural populations of mosquitoes in the *Anopheles gambiae* complex. Paracentric chromosomal inversions have been studied extensively in the *A. gambiae* complex and are shown to correlate

Methods

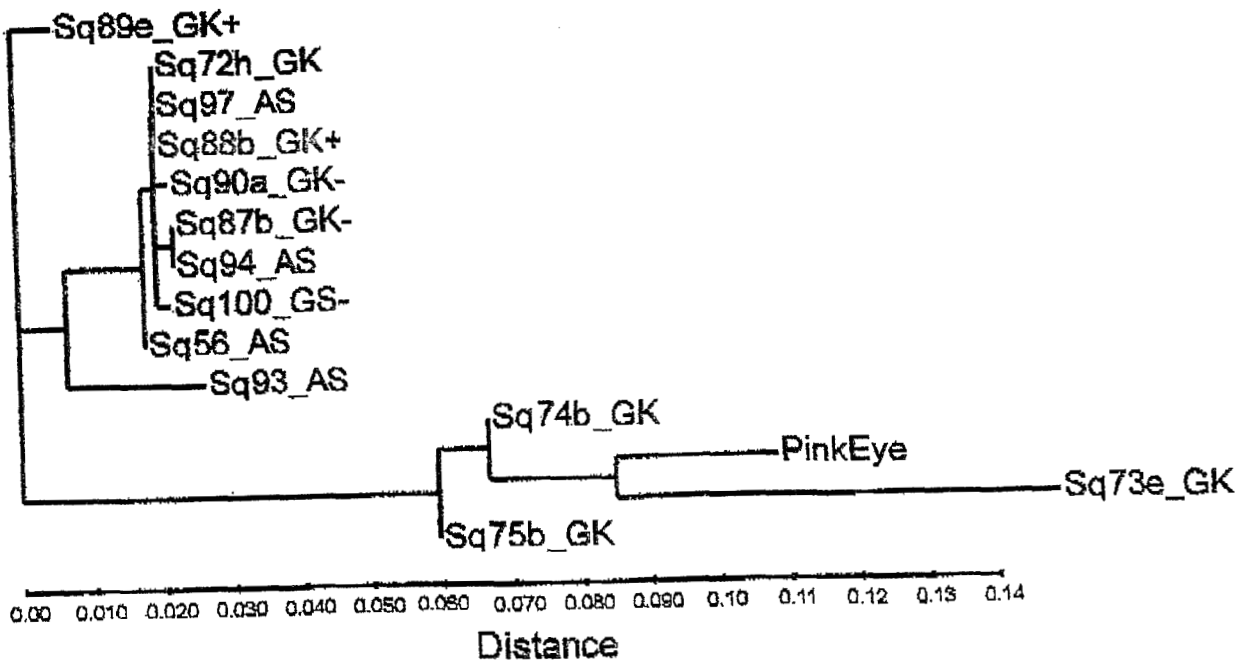
Our samples include mosquitoes collected in Senegal at two localities about 200 km apart, one of which is a study site for hyperendemic malaria (Trape et al, 1994). The Kenyan collections are all from within 30 km of Kisumu, also including a study site for hyperendemic malaria (Githeko et al., 1992). We are sequencing DNA at several genetic loci of known location on the physical map (data provided by Dr. Frank Collins). We have designed sets of nested primers to allow amplification from minute samples using nested PCR. This allows the same mosquitoes to be used for other assays, such as karyotyping, sporozoite ELISA, blood meal ELISA, and other DNA analyses (e.g., with micro-satellite markers). Karyotype analysis is being performed by Ousmane Faye, University of Senegal, Dakar, and by Odette Mukabayire, Malaria Branch, Division of Parasitic Diseases, CDC, Atlanta, both of whom have received training from Dr. Mario Coluzzi.

PCR is performed using *Pfu* DNA polymerase to reduce mis-incorporations. Sequencing is performed by the University of Arizona Macromolecular Structure Facility on an Applied Biosystems model 373a automated sequencer using fluorescent dye chain terminators. PCR products were either sequenced directly or after cloning into pCR-Script (Stratagene). Cloning was required for heterozygotes for major differences, especially insertion-deletion events. So far we have generated data for three variable loci from two different inversions described in Table 1.

Our conceptual approach to analysis of these data is to interpret sequence variants as a pattern of superimposed mutations, or synapomorphies. Estimating population-genetic parameters such as effective population size and gene flow from such data requires two conditions: 1) that one can accurately interpret the pattern of synapomorphies, and 2) that the sequences under study are not under unusual selective forces (purifying or balancing). Our first approximation to estimating the history of mutational changes is to use phylogenetic trees for each locus which are presented here as Figures 1-3. Because the nature of selection at our loci is impossible to know objectively, we restrict ourselves to comparisons within loci, attributing differences in genetic diversity between loci to selection. In the future will apply analyses which accommodate recombination within loci as well as among them.

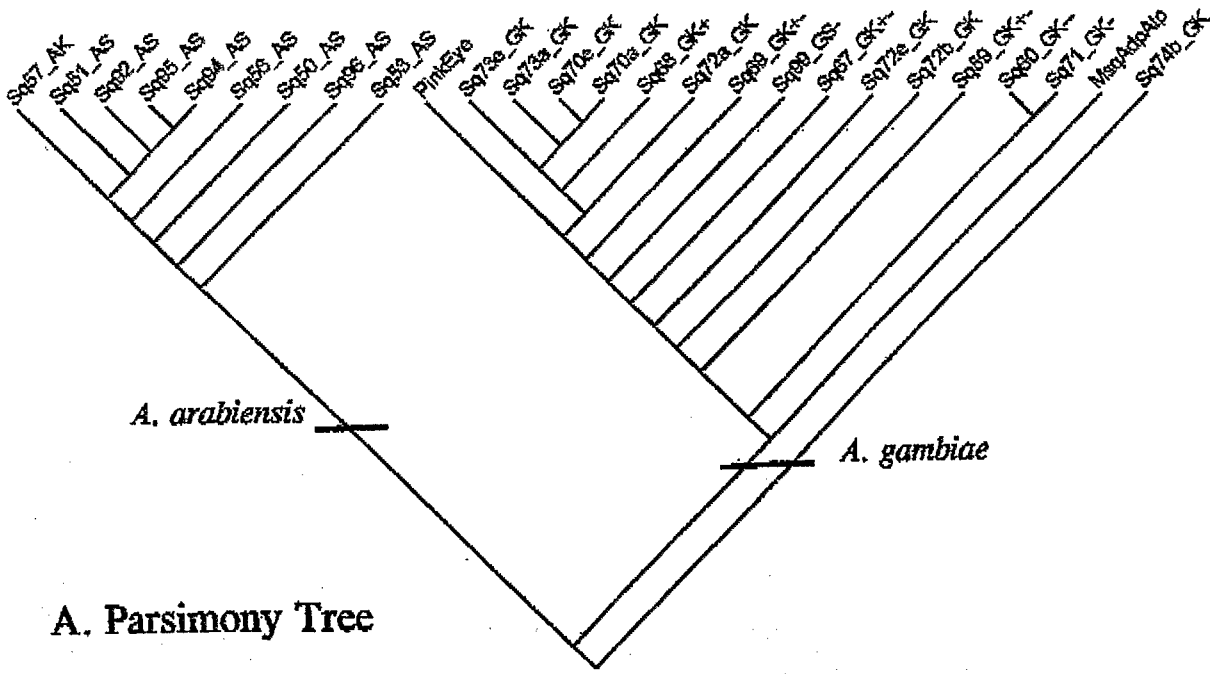


A. Parsimony Tree

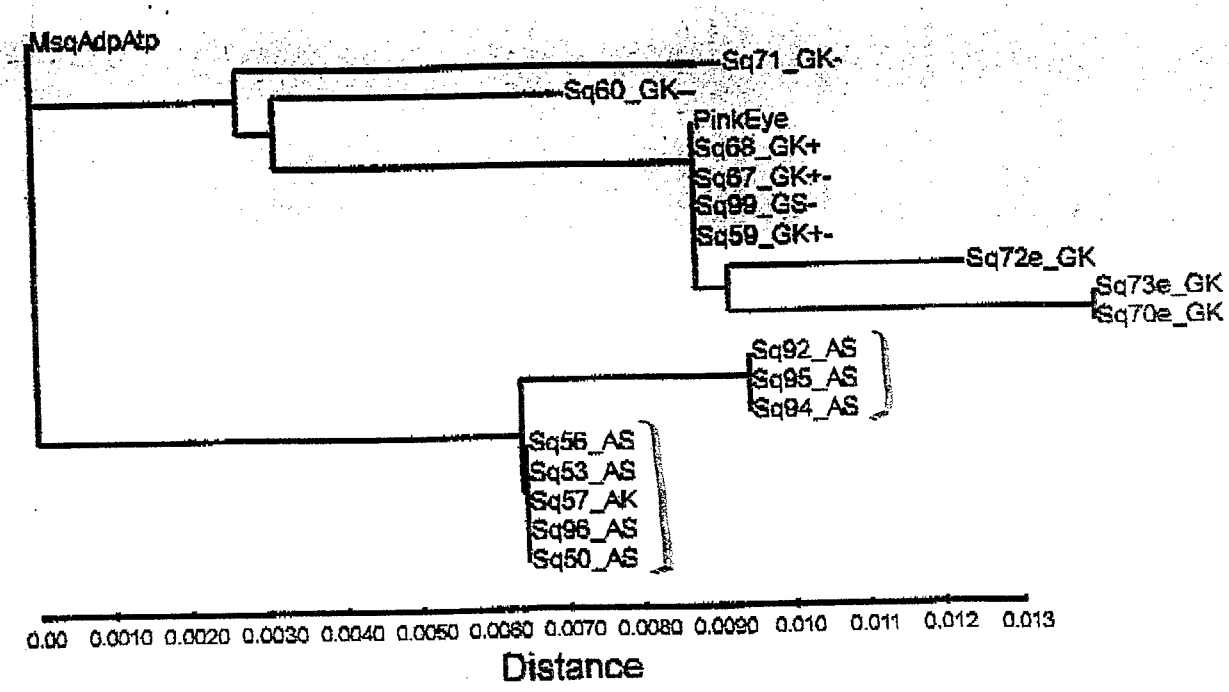


B. Least-squares Distance Tree

Figure 1. Phylogenetic trees of sequence variants at cDNA76 (Transcription Initiation Factor) in inversion 2La. We present both maximum parsimony (A.) and distance analyses (B.). Names refer to individual mosquitoes or clones with indications of species, locality, and inversion type: a "G" indicting *Anopheles gambiae* and "A" indicating *A. arabiensis*; "S" means Senegal, "K" means Kenya. "+" means 2La-Standard, "-" means 2La-Inverted. There is a strong tendency for mixing of all categories on this tree,

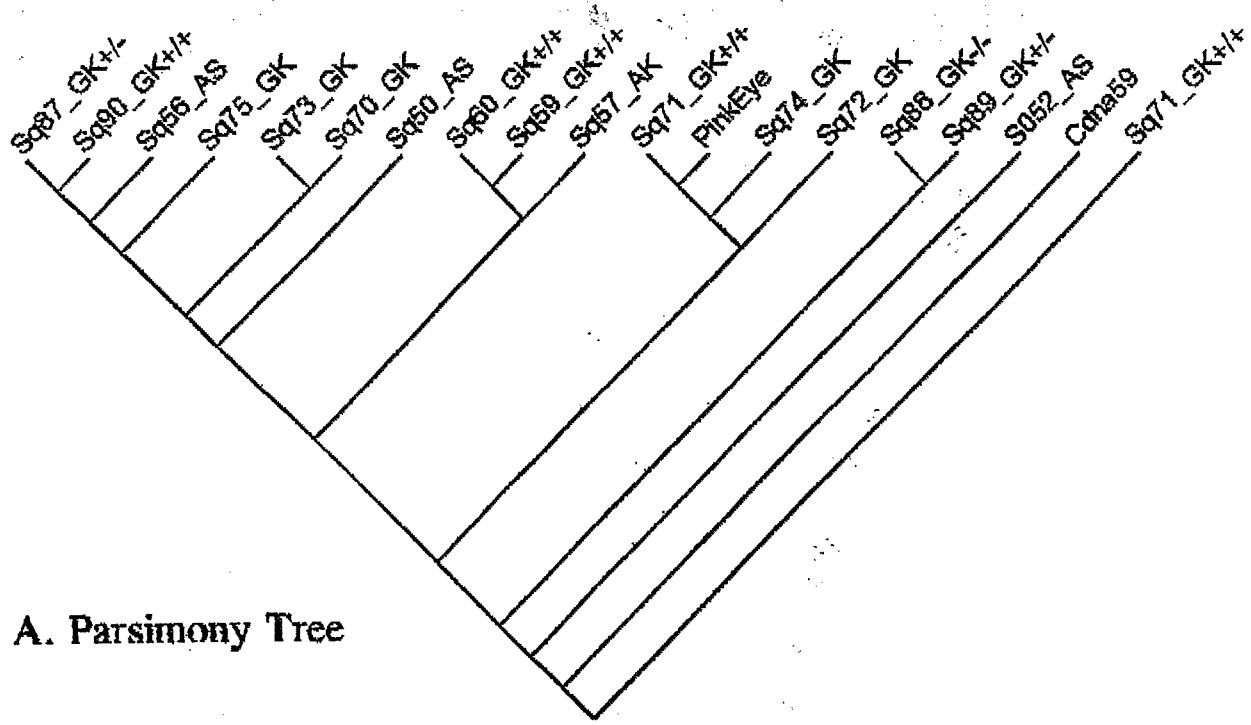


A. Parsimony Tree



B. Least-squares Distance Tree

[redacted] sequence variants at Adp/Atp translocase in inversion 2La. At this locus,



A. Parsimony Tree

Sq71_GK+/+

Preliminary Interpretations

Although we are in the middle of data-collection, with many individuals of known karyotype still to add to our trees, we believe there are some interpretations which can be derived from the data in hand. The questions framed at the outset of this work seem to have the following tentative answers, subject to change as more data become available:

1) Is variation of DNA sequences in regions affected by inversions structured strongly by inversion orientation? The answer seems to be "yes and no." At the Adp/Atp locus, sequences known to represent the Standard orientation of 2La form a cluster (also including at least one Inverted) to the exclusion of other Inverteds and all of *A. arabiensis* (which is fixed for Inverted at this inversion). This is consistent with Inverted being the primitive condition and the inversion generating Standard being a recent event, with diversity among Standard sequences arising subsequent from that time. If further data bear this up, it could signal a strong effect of inversions on genetic structure, particularly at loci near the inversion break-points, as is the case for Adp/Atp.

2) Is variation structured strongly by classification into the species *gambiae* and *arabiensis*? The tentative answer is "yes and no." At one of three loci, Adp/Atp translocase, is there a strong separation of the two

Literature Cited

Besansky, N. J., J. R. Powell, A. Caccone, D. M. Hamm, J. A. Scott, and F. H. Collins. 1994. Molecular phylogeny of the *Anopheles* gambiae complex suggests genetic introgression between principal malaria vectors. Proc. Natl. Acad. Sci. USA 91: 6885-6888.

Coluzzi, M., A. Sabatini, V. Petrarca, and M. A. Di Deco. 1979. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Trans. Roy. Soc. Trop. Med. Hyg., 73: 483-497.

Githeko, A. K., A. D. Bradling-Bennet, M. Beier, F. Atieli, M. Owaga, and F. H. Collins. 1992. The reservoir of *Plasmodium falciparum* malaria in a holoendemic area of western Kenya. Trans. R. Soc. Trop.