

## **CONCLUSION ON THE TREE.**

There were five different groups of sample sequences from the alignment file each with a distinct capital letter(s) at the beginning representing the geographical location from where the sample was collected followed by a number and the groups distinct letter were E, K, T, U, UK.

A phylogenetic tree was inferred using a maximum likelihood algorithm from an alignment file (SNPS\_alg.fasta) comprises of 236 sample sequences from all the five groups and for assessing confidence of the branches of the tree 1000 bootstrap were performed.

From the tree file (SNPS\_alg.fasta.treefile) most of the branches were having bootstrap value above 70% while some of the few branches were having bootstrap value lower than 70% which indicate lower degree of confidence which may rise uncertainty on their reliability during interpretation of the tree.

In most of the part of the tree, sequences from the same geographical areas tend to form clusters which may indicate relatively equal evolutionary distance from the common ancestor, also showed at some part sample sequences from different areas cluster together this may indicates that the variation inferred from the tree is also base on genotype of an individual rather than the geographical area of collection. Thus its better to explain the diversity of these species in terms of genotypic groups inferred by the tree rather than basing on their geographical origins.

When rooting the tree from the midpoint using Figtree software, we obtain two branches which serves as internal nodes for other branches on the tree. In both of the branches there are mixtures of sequence clusters but also shows some dominant sample sequences from each respective internal nodes which indicates both the effect of environment and mutation in producing variation among species.