- 0. Given a gene-by-array expression profile matrix, M_{g^*p} , with g genes and p arrays. The following process can be applied for either gene or array grouping and sorting.
- 1. Compute the gene-by-gene proximity (e.g., correlation, Euclidean distance, etc.) matrix R_{g^*g} , (Figure 3a).
- 2. Select an HCT linkage type (e.g., single, complete, average, centroid, etc.) and construct the hierarchical clustering tree, HCT, for R_{g^*g} with g terminal nodes, T_i (*i*=1, ..., g), and g-1 intermediate nodes, N_j (*j*=1, ..., g-1). We further define N_j^{T} (*j*=1, ..., g-1) the set of terminal nodes spanned by intermediate node N_j , (Figure 3b).
- 3. Apply the converging sequence of iteratively formed correlation matrices [1] to R_{g^*g} and obtain the rank-two ellipse seriation, R2E, for R_{g^*g} . Define relative positions (1, ..., g) of the g terminal nodes, T_i (i=1, ..., g), in R2E as O_i (i=1, ..., g) and define N_j^O (j=1, ..., g-1) the set of relative positions of terminal nodes spanned by intermediate node N_j , (Figures 3c and 3b).
- 4. Use *R2E* as an external reference to guide the flipping mechanism of *HCT* to obtain the final flipped tree, *HCT_R2E*. For an intermediate node N_j with two sub-nodes N_{j1} and N_{j2} , assign N_{jk} as the upper sub-node of N_j if mean (N_{jk}^{O}) =min $(mean(N_{j1}^{O}), mean(N_{j2}^{O}))$, (Figure 3d).
- 5. Sort M_{g^*p} according to the relative terminal node positions of HCT_R2E , (Figure 3d).