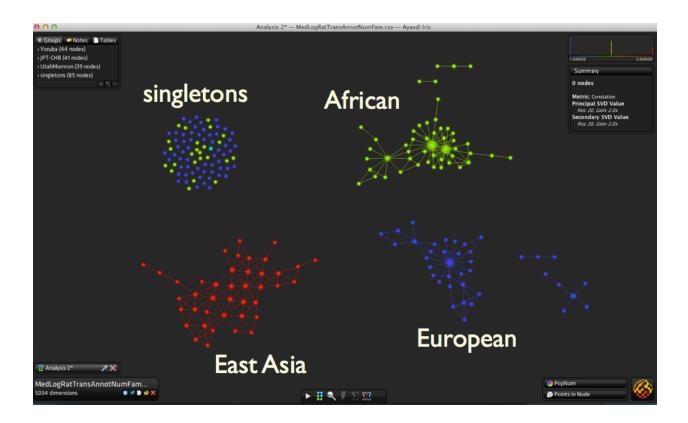
Analyzing Population Structure using CNV

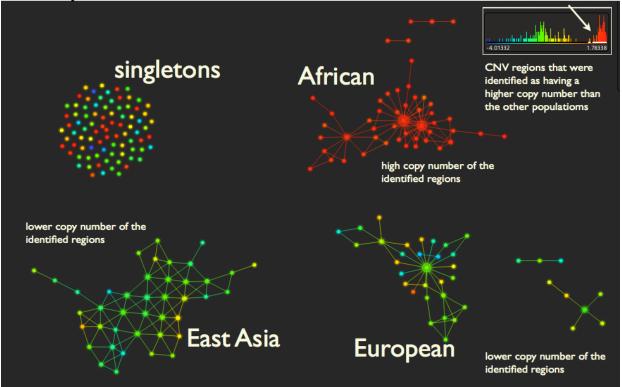
In this example, we analyzed the population structure using copy number variation data from the HapMap samples (PMID 19812545). There were 313 samples and over 5000 CNV (copy number variations) regions. The samples comprised of European, African or East Asian ancestry. A dot text file containing samples as rows and CNV regions as columns were fed into Iris and an automated topological analysis was performed on the data set. Using a primary and secondary PCA lens and a correlation metric, we were able to separate samples according to their ancestry using CNV regions into discrete networks.



To understand what CNV regions best discriminate these population networks, we can use our statistical feature Explain to identify those CNV regions. CNV regions that differentiate the population of African, European and East Asian ancestries are found in the Tables 1a,b and c. Using the feature Data Field (Average), we colored the nodes by the values of those CNV regions identified as being high in the African ancestry.

Copyright 2013 Ayasdi Inc.

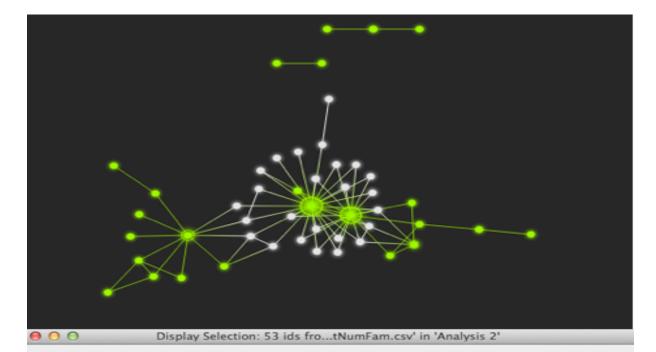
Nodes are colored by the average values of those CNV regions identified as significantly higher for the population of African ancestry



In addition to the separation of the ancestry, we also saw interesting "textures" that were associated with the underlying data. As seen in figure 2 below, the African and European ancestry networks had a "starburst" structure whereas the East Asian ancestry network had no such structure. Upon a closer inspection of the starburst structure by using the Display feature, it is evident that the "starburst" structure comprised of parent-child relationships, that is, the samples in each of the node were either a father-child or a mother-child or in some cases, father-mother-child samples (Table below figure 2). Through this example, we show that topological analysis was able to not only identify discrete population structure easily but also identified fine textures of the data, including parent-child relationships. In addition, the regions that are important in determining the population structure are easily identifiable through our Explain feature.

Copyright 2013 Ayasdi Inc.

Discover what you don't know.



Pactorvalu	FactorValueSex	Sample	FamilyAll	Relationship	Family
YORUBA	male	NA19249	Y120-01	child	Y120
YORUBA	male	NA19248	Y120-03	father	Y120
YORUBA	female	NA19238	Y117-02	mother	Y117
YORUBA	female	NA19237	Y116-01	child	Y116
YORUBA	female	NA19235	Y116-02	mother	Y116
YORUBA	male	NA19236	Y116-03	father	Y116
YORUBA	male	NA19194	Y112-01	child	Y112
YORUBA	female	NA19193	Y112-02	mother	Y112
YORUBA	male	NA19192	Y112-03	father	Y112
YORUBA	male	NA19191	Y111-01	child	Y111
YORUBA	female	NA19190	Y111-02	mother	Y111
YORUBA	male	NA19174	Y100-01	child	Y100
YORUBA	female	NA19118	Y100-02	mother	Y100
YORUBA	male	NA19117	Y100-03	father	Y100
YORUBA	female	NA19115	Y079-01	child	Y079
YORUBA	female	NA19114	Y079-02	mother	Y079
YORUBA	female	NA19129	Y077-01	child	Y077
YORUBA	female	NA19127	Y077-02	mother	Y077
YORUBA	female	NA19148	Y075-01	child	Y075
YORUBA	female	NA19147	Y075-02	mother	Y075
YORUBA	male	NA19123	Y061-01	child	Y061
YORUBA	female	NA19122	Y061-02	mother	Y061
YORUBA	male	NA19121	Y061-03	father	Y061
YORUBA	male	NA19120	Y060-01	child	Y060

Copyright 2013 Ayasdi Inc.

New List

Close Table

Save Table