Step 1: Accurate definition of the copy number profile in each sample

Individual .CEL Files

Array Calibration
Copy Number Estimation
Segmentation

Segmented Copy Number Profiles

Step 2: Identification/separation of underlying SCNAs

- Elimination of arm-level SCNAs by use of amplitude threshold
- Deconstruction of segmented profile into underlying SCNAs
  Allows for modelling of background rate of SCNAs and length-based separation of arm-level and focal SCNAs

Step 3: Scoring SCNAs in each region according to likelihood of occurring by chance

- G = frequency x amplitude
  p-values computed by random permutation of markers across genome
- G = -log(Probability | Background)
  Scores computed on markers or genes
  p-values computed by random permutation of markers or bins across genome

Step 4: Defining independent genomic regions undergoing significant levels of SCNA

- Greedy segment peel-off algorithm
  Iteratively subtracts segments covering each peak and rescores until no significant peaks remain on chromosome
- Arbitrated peel-off algorithm
  Formalizes idea that segments can have multiple targets by allowing segment scores to be split among multiple potential peaks during peel-off

Step 5: Accurate definition of the copy number profile in each sample

- Leave-k-out
  Assumes that at most ÔkÕ passenger events aberrantly define the minimal common region
- RegBounder
  Models expected local variation in G-score to define boundaries predicted to contain the true target with predetermined confidence

GISTIC 1.0 (Beroukhim et al, 2007)
GISTIC 2.0