IMPROVED STR ARTIFACT RECOGNITION IN MIXTURES USING OSIRIS Robert Goor, George Riley, Douglas Hoffman, Stephen Sherry, National Center for Biotechnology Information, National Library of Medicine, NIH

Recently, laboratories have begun to use analytical thresholds below 30 RFU to analyze complex mixed profiles, and so it has become critical for analysis software to be able to distinguish low level artifacts from actual alleles. OSIRIS employs mathematical modeling to achieve this. One of the most difficult artifacts to analyze is pull-up because it is virtually impossible to analyze based solely on single instances. How can the software distinguish between the two possible scenarios – peaks that are truly pull-up and peaks that simply comigrate by coincidence? The latter is particularly important in mixture analysis, because a minor peak that is an actual allele may also appear to be pull-up.

Because pull-up is caused by a matrix mismatch between the channels, pull-up manifests as a sample-wide pattern; peaks with equivalent amounts of DNA in the same channel will cause equivalent pull-up in another channel. So it is only at the sample-wide level that pull-up artifacts can be conclusively identified. Furthermore, any analysis of pull-up must be able to distinguish the two scenarios above. The latest version of OSIRIS now includes a sample-wide mathematical analysis. Using peak height and coincidence data across pairs of channels, OSIRIS uses a mathematical tool called least median of squares to determine what pattern may exist and which peaks are outliers from the pull-up pattern. An outlier, in this analysis, does not match the characteristics of a pure pull-up peak and therefore must be an allele. Removing the outlier (allele) peaks from pull-up consideration allows OSIRIS to use ordinary regression to determine the true pull-up pattern. Very large peaks where the data is off-scale are considered separately because they exhibit a different pattern from the peaks that are not off-scale.

The result is an accurate assessment of which peaks are pull-up, and receive no allele call, and which are alleles, with perhaps some pull-up correction in the case of coincident alleles. For analysts, this translates into a more accurate determination of which small peaks are pull-up, reducing the editing burden and freeing up time for case interpretation.