CPPs in Strand Displacement Amplification Assay

Background

Strand displacement amplification (SDA) is an isothermal process that permits 10^{10}-fold amplification of a DNA target sequence in as little as 15 min. SDA is used to assay nucleic acids for routine application in the clinical laboratory. The isothermal nature of the reaction process offers distinct advantages with regard to the cost and simplicity of instrumentation, while a universal detection format permits the use of the same fluorescent detector probes across multiple analytes. This has important potential in the field of genetic analysis, in which disease predisposition and therapeutic efficacy are frequently determined by multiple nucleic acid markers.

Objective: Address Buffer and Salt Performance Issues

SDA reactions require appropriate concentrations of buffers and salts to achieve target assay performance. For delivery of practical clinical assays, it was deemed necessary to identify a single buffer combination that could be used with multiple assay types. This will ensure a sample processing workflow that is simple and easy to follow.

To achieve these objectives, it was decided to identify the critical process parameters related to buffers and salts and use the analysis to identify a viable buffer/salt design space.

Performance Data: Summary

The BD data frame is comprised of 89 cases (rows) each with 5 dimensions (columns). The measure (response) is the DNA amplification rate—that is, the exponential amplification rate on a log scale (lnAmp).

```
attach(BD)
summary(BD)
```

Performance Data: Experimental Space Coverage

There are 89 runs with four buffer components (KCL, MgCl2, KiP04, dnTP) systematically varied in a space-filling design. A two-dimensional (hexagonal binning) scatterplot displays the coverage of the experimental space.

```
hexplom(
   BD,
   colramp=BTY,
   varname.cex=0.9,
   axis.text.cex=0.7,
   # xbins determines number of buckets hor and vert
   # affects the visual size of the hexagons
   xbins=15
   )
```

The scatterplot is comprised of a 15 x 15 matrix of hexagon bins containing a count of the number of runs that fall within the boundaries of the bin. Count values are displayed as a color gradient running from yellow (low count) to blue (high count).

Critical Process Parameters: Identification and Prioritization

A Random Forest regression analysis of the data based on a model comprised of an ensemble of decision trees. The analysis shows the fraction of the total variance explained by the model—high % indicates confident analysis; lower % indicates a lot of noise in the data.

```
dims <- BD[ ,1:4]
rf <- randomForest(x=dims, y=lnya, mtry=2)
rf
```

A first powerful feature of Random Forest regression is that it provides a prioritized list of variables with regards to how critical they are to performance prediction—that is, to what extent are they critical process variables.

```
barchart(
   ~ sort(rf$importance[,1]),
   aspect=0.5,
   xlim=c(0,140),
   xlab="Importance to DNA Amplification Rate"
   )
```

Critical Process Parameters: Extent of Criticality

A second powerful feature of Random Forest regression is that it provides a graphical depiction of the marginal effect of a given process parameter (the derivative; slope) on the assay performance measure—that is, the amplification factor—while averaging out the effect of all the other parameters. Highly critical process parameters exhibit steep slopes in partial dependence plots.

```
par(mfrow = c(2,2))
partialPlot (rf, pred.data=dims, x.var=KCl, ylim=c(8,11))
partialPlot (rf, pred.data=dims, x.var=MgCl2, ylim=c(8,11))
partialPlot (rf, pred.data=dims, x.var=KPO4, ylim=c(8,11))
partialPlot (rf, pred.data=dims, x.var=dNTP, ylim=c(8,11))
par(mfrow = c(1,1))
```

Assay Performance: Visualization

Once we have identified the critical process parameters, we can visualize their impact on assay process performance by fitting the collection or irregular data to a thin plate spline surface. We then use the resulting model to calculate a contour plot for assay process performance.

```
# fit surface to data
fit <- Tps(x=dims, Y=lnya)

# surface of variables "x" and "y" holding other two fixed at their median levels

out.21 <- predict.surface(fit, xy=c(2,1))
out.31 <- predict.surface(fit, xy=c(3,1))
out.41 <- predict.surface(fit, xy=c(4,1))
out.32 <- predict.surface(fit, xy=c(3,2))
out.42 <- predict.surface(fit, xy=c(4,2))
out.43 <- predict.surface(fit, xy=c(4,3))

# contour levels top be displayed
contLev <- c(6,7,8,9,10,11,12,13)

# surface plot
par(mfrow = c(2,3))
plot.surface(out.21, type="C", zlim=c(5,14),levels=contLev)
plot.surface(out.31, type="C", zlim=c(5,14),levels=contLev)
plot.surface(out.41, type="C", zlim=c(5,14),levels=contLev)
plot.surface(out.32, type="C", zlim=c(5,14),levels=contLev)
plot.surface(out.42, type="C", zlim=c(5,14),levels=contLev)
plot.surface(out.43, type="C", zlim=c(5,14),levels=contLev)
par(mfrow =c(1,1))
```

Assay Performance: Uncertainty in Visualization

We obtain an estimate of the uncertainty in the assay performance visualization by calculating the standard error on the surface and plotting it as a contour plot.

```
# calculate and display as a surface the standard error
out.32.se <- predict.surface.se(fit, xy=c(3,2))
plot.surface(
   out.32.se,
   type="C",
   xlab="MgCl2",
   ylab="KiPO4"
   )
```

Analysis Conclusions: Comparison to Reported by Walker

The results of the above analysis can be compared to the published results of Walker, who noted the following:

- "SDA reactions usually contain 50 mM KiPO4 (pH 7.4) which serves as a buffer and K+ source for the enzymes. In preparing this buffer, one must be careful not to raise the K + concentration any higher than necessary while adjusting the pH because K+ concentrations >100 mM are detrimental to SDA." [We observe this in partial dependence plots].
- "We mix 13 ml of 0.5 M KH2PO 4 with 87 ml of 0.5 M K2HPO 4 to form 0.5 M KiPO 4 (pH 7.6) (101 stock solution), which corresponds to a final lx concentration of 94 mM K+ in the SDA reaction." [Optimal concentration are consistent with our analysis]
- "SDA reactions usually contain 6 mM MgC12 and 4 mM total dNTP.. However, because the dNTP and free Mg z+ concentrations are coupled, the two concentrations cannot be changed independently." [Coupling (interaction) between MgCl2 and dNTP is observed in contour plots of assay performance visualization]
- "The free Mg 2+ concentration is ~2 mM because dNTPs quantitatively bind Mg z+. Other concentration combinations of MgC12 and dNTPs may produce favorable results." [Consistent with the multivariate nature of the performance of the assay process.

References

Walker, G.Terrance., "Empirical aspects of strand displacement amplification," *PCR Methods Appl.* 1993 (3) 1-6

Hellyer, Tobin J. and Nadeau, James G., "Strand displacement amplification: a versatile tool for molecular diagnostics," Expert View of Molecular Diagnostics, 2004 (Vol 4 No 2) 251-261.

Data from: Becton Dickinson Research Center, Research Triangle Park, NC