

Principle behind Cross Hybridization adjustment mathematics:

Example:

$$\text{Total Sig 1} = \text{True Sig 1} + \text{CH adj} \times \text{True Sig 2} + \text{CH adj} \times \text{True Sig 3}$$

$$\text{Total Sig 2} = \text{True Sig 2} + \text{CH adj} \times \text{True Sig 1} + \text{CH adj} \times \text{True Sig 3}$$

$$\text{Total Sig 3} = \text{True Sig 3} + \text{CH adj} \times \text{True Sig 1} + \text{CH adj} \times \text{True Sig 2}$$

Total Sig – total signal on the probeset

True Sig – signal due to correct fragment binding

CH adj – CH adjustment factor specific to each probeset

Estimating the value of CH adj from validation experiments. Known quantities are Total Sig 1, 2, and 3, and True Sig 1, 2, and 3 (derived from different experiments where only one species gDNA is hybridized. Total signal for that phylospecies probeset is treated to be equal to true signal). Thus each equation above has only one unknown, CH adj, and gives a separate estimate of CH adj. These estimates are averaged.

Estimating True Signals. Known quantities are Total Sig 1, 2, and 3, and CH adj factors. Unknown quantities are True Sig 1, 2, and 3. This leaves 3 equations in 3 unknowns, which can be solved mathematically.

In the experiment with the Microbiota Array, instead of 3 signals (1, 2, and 3) there are 775 signals, but the principle is the same. Note that CH adjustment is set to zero for cross-hybridization from the probesets/species outside the same genus. Thus the CH adjustment for the contribution of True Sig 3 to Total Sig 1 would be set to zero if signals 1 and 3 were from different genera.