

### ***Alternative Simulation Model***

To make sure that our results and conclusions were not an artifact of our data simulation model, a second set of simulated data was generated using the model for high-rearrangement profiles as described by Hupé et al. without outlier addition (Hupe, et al., 2004). Again, results were based on 500 samples with 20 chromosomes each consisting of 100 clones.

Generally, this data was less noisy than the data simulated using our model (Figure S 4) and merging only provided large improvement for proportion of the discordant pairs measure leaving three other measures largely unchanged (See supplementary Table S 3, Table S 4 and Table S 5). Moreover, likely due to the low experimental variability, merging was able to improve specificity without decreasing sensitivity. By just accepting a localization error of one clone ( $w=1$ ), both HMM and DNACopy were able to identify all breakpoints. At the same time, DNACopy was very specific and did not predict any excessive breakpoints (See supplementary Figure S 5). Both GLAD and HMM had a problem detecting the exact change points. For HMM, this may be due to the model on which the HMM is based, where the possibility of changing states may be too small or depending too much on preceding or following clones. Overall, GLAD was less sensitive than the other methods, hence, sensitivity did not increase as much when increasing the accepted localization error.

### ***References***

Hupe, P., Stransky, N., Thiery, J.P., Radvanyi, F. and Barillot, E. (2004) Analysis of array CGH data: from signal ratio to gain and loss of DNA regions, *Bioinformatics*, **20**, 3413-3422.