

Figure S2: Identification of whole and partial gene deletions/duplications by coverage analysis. Relative coverage analysis was routinely performed in groups of 12 samples from each experiment. At each target base, the relative coverage was calculated by dividing the observed depth of coverage by the average coverage of all targeted bases in a given sample; this value was then normalized to the average relative coverage at each base for all samples in the analysis cohort. Thus for normal copy number, a normalized relative coverage of 1 would be expected, while samples containing heterozygous deletions or duplications would show values of 0.5 or 1.5 respectively across the affected regions. Calculations were performed in STATA and data visualized for each gene of interest in MS Excel. Normalized relative coverage for each sample was plotted against base position; heterozygous deletions/duplications were identified by visual inspection for extended regions which occurred outside two standard deviations from the mean normalized relative coverage and which had values close to 0.5 or 1.5 respectively. Figures S2a-e show relative coverage analysis plots revealing whole and partial gene deletions, and partial gene duplications, in the genes indicated. In each plot, the black line indicates a positive control sample with a known deletion/duplication previously identified by MLPA or aCGH, corresponding to the samples listed in table 1; grey shading indicates  $\pm 2$  standard deviations from the mean normalized relative coverage at each position.

