Figure 3: Comparison of BH(○), AP(△), localFDR(×), OR (+), and LIS (□) in basic simulation 2: (1) chain-MRF, (2) tree-MRF, (3) grid-MRF; (a) FDR vs ∆(φ), (b) FNR vs ∆(φ), (c) ATP vs ∆(φ), (d) FDR vs ∆(μ), (e) FNR vs ∆(μ), (f) ATP vs ∆(μ).

are still valid, but the data-driven procedure becomes more and more conservative as we increase the variance of μ. The data-driven procedure can be more conservative than the BH procedure when ∆(μ) is large enough. The conservativeness appears most severe in the grid-structure. However when we look at the FNR and ATP, the data-driven procedure still dominates BH, AP and localFDR substantially in all the situations, although the data-driven procedure loses a certain amount of efficiency compared with the oracle procedure when the variance of μ gets large.

Figure 4: Comparing BH(○), AP(△), localFDR(×), OR (+), and LIS(□) in basic simulation 3: (a) FDR vs n, (b) FNR vs n, (c) ATP vs n.

Figure 4 shows the results from basic simulation 3. The oracle procedure and localFDR are liberal when the sample size is small. This is because when the sample size is small, there exists a discrepancy between the true distribution of the test statistic and the limiting distribution. Quite surprisingly, the data-driven procedure stays valid. The reason is that the data-driven procedure can estimate the parameters from data. The data-driven procedure and the oracle procedure still have comparable performance and enjoy a much lower level of FNR compared with the baselines. For all the basic simulations, we set the nominal FDR level to be 0.10. We have also replicated the basic simulations by setting the nominal level to be 0.05, and similar conclusions can be made.

4 Simulations on Genetic Data

Unlike the fabricated dependence structures in the basic simulations in Section 3, the dependence structure in the simulations on genetic data in this section is real. We simulate the linkage disequilibrium structure of a segment on human chromosome 22, and treat a test of whether a SNP is associated as one individual test. We follow the simulation settings in the work of Wu et al. (2010). We use HAPGEN2 (Su et al., 2011) and the CEU sample of HapMap (The International HapMap Consortium, 2003) (Release 22) to generate SNP genotype data at each of the 2,420 loci between bp 14431347 and bp 17999745 on Chromosome 22. A total of 685 out of 2,420 SNPs can be genotyped with the Affymetrix 6.0 array. These are the typed SNPs that we use for our simulations. Within the overall 2,420 SNPs, we randomly select 10 SNPs to be the causal SNPs. All the SNPs on the Affymetrix 6.0 array whose r^2 values, according to HapMap, with any of the causal SNPs are above t are set to be the associated SNPs. In the simulations, we report results for three different t values, namely 0.8, 0.5 and 0.25. We also simulate three different genetic models (additive model, dominant model, and recessive model) with different levels of relative risk (1.2 and 1.3). In