



Figure S.2: Map configuration in the four simulated scenarios. From top to bottom, left to right: SIM1, SIM2, SIM3 and SIM4. SIM1,  $q = 100$  transcripts simulated with SNP 30 and 239 influencing transcripts 1-20 and 71-80, SNP 161 influencing transcripts 17-20, SNP 225 influencing transcripts 91-100, and finally eQTLs 362 and 466 influencing transcripts 81-90. Altogether 94 transcript-SNP associations are simulated in 50 distinct transcripts; SIM2, 100 responses simulated with only three hot spots (30, 161, 239) and the same simulated pattern of association as in the first scenario leading to 64 transcript-SNP associations in 30 distinct transcripts; SIM3, the simulation set-up is identical to the first scenario for the first 100 responses, but the number of simulated responses is increased to  $q = 1,000$ , simulating further 900 transcripts from the noise; SIM4, as in the second simulated data set for the first 100 responses, with additional 900 responses simulated from the noise, and altogether  $q = 1,000$ . The symbol ‘G’ in the  $y$ -axis identifies groups of transcripts that are influenced by the same pattern of markers. SIM1 and SIM2,  $G_1$ : transcripts 1-16;  $G_2$ : transcripts 17-20;  $G_3$ : 21-70;  $G_4$ : transcripts 71-80;  $G_5$ : transcripts 81-90;  $G_6$ : transcripts 91-100. SIM3 and SIM4 as before with  $G_7$ : transcripts 101-1000.