Supplementary Figures

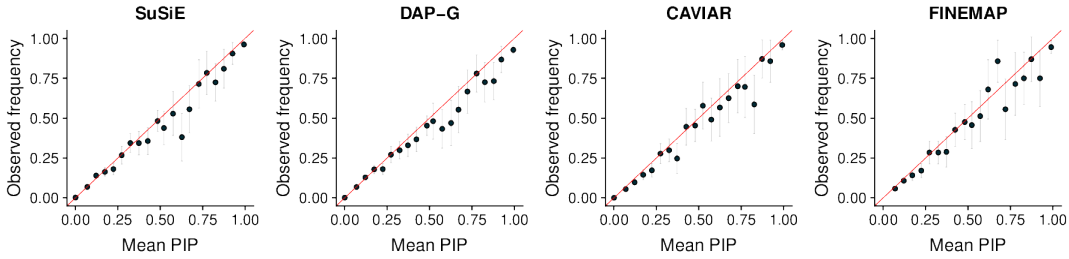

FIGURE S1 Assessment of PIP calibration. Variables across all simulations were grouped into bins according to their reported PIP (using 20 equally spaced bins, from 0 to 1). The plots show the average PIP for each bin against the proportion of effect variables in that bin. A well calibrated method should produce points near the $x = y$ line (the diagonal red lines). Gray error bars show ± 2 standard errors.

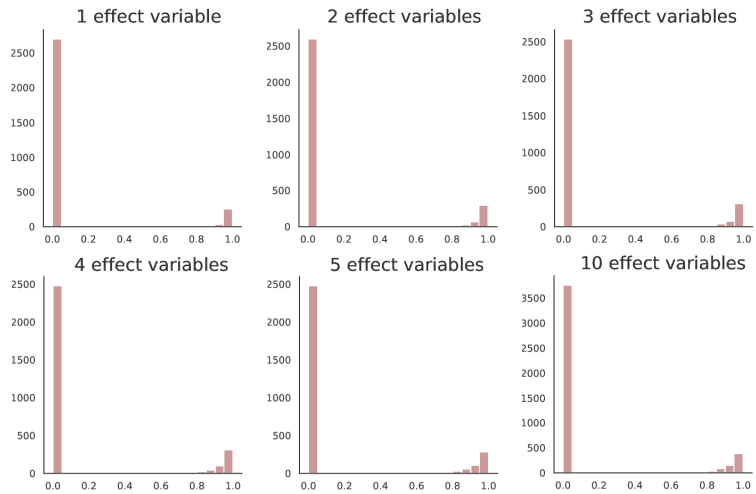

FIGURE S2 Distribution of purity for 95% credible sets for different numbers of effect variables. Histograms for 1–5 effect variables are obtained from all 95% credible sets produced by SuSiE in the first simulation scenario, with $S = 1, \ldots, 5$, as described in Section 4 of the main text, and the 10 effect variables histogram is obtained from all 95% credible sets produced by SuSiE in the second simulation scenario, with $S = 10$.

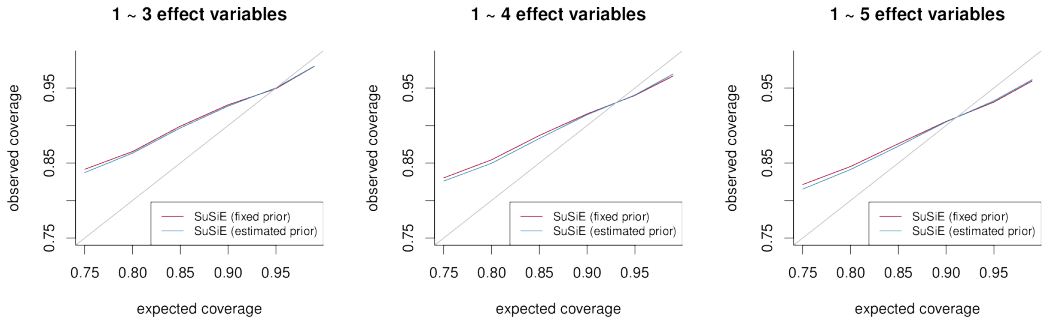

FIGURE S3 Additional assessment of SuSiE CS coverage. These three plots show coverage of SuSiE credible sets as ρ (the probability that the credible set contains at least one effect variable; see Definition 1 in the main text) is varied from 75% to 99%. Proportions shown in the vertical axis are based on all credible sets generated by SuSiE in simulations from simulation scenario 1, with different simulation settings for S, the number of effect variables. Consistent with Fig. 3, coverage decreases with the inclusion of weaker signals.

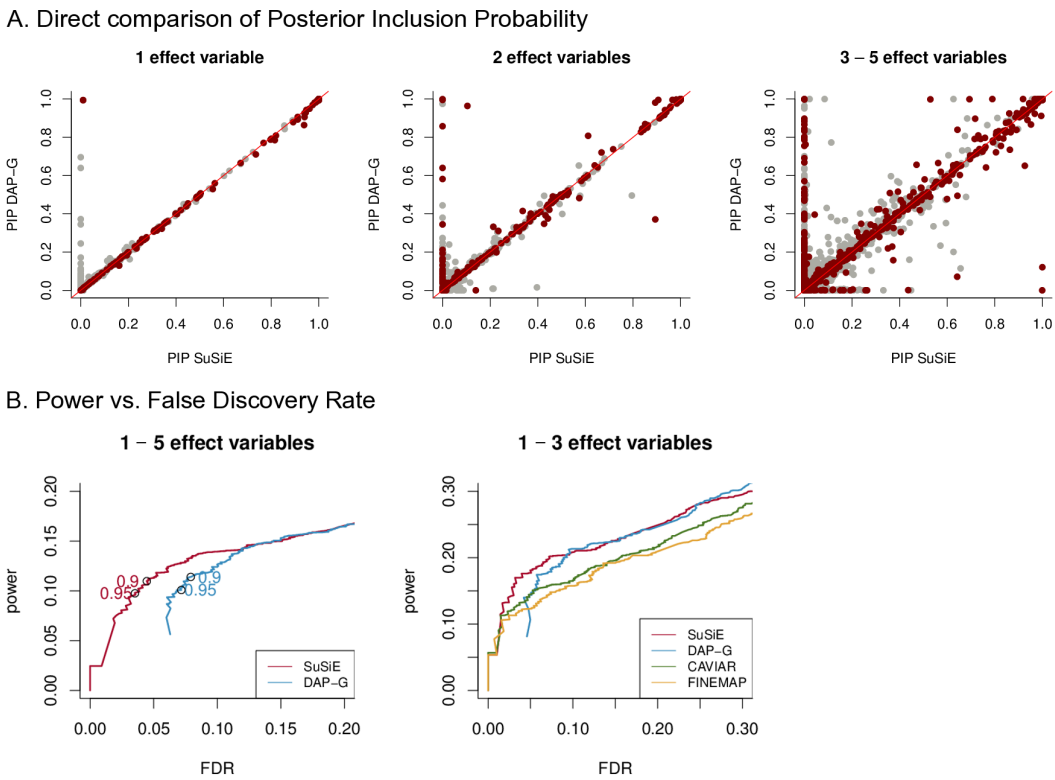

FIGURE S4 Comparison of posterior inclusion probabilities (PIPs) computed by SuSiE, in which the prior variances σ^2 are estimated rather than fixed to 0.1, against PIPs computed by DAP-G, and by other methods. The results shown here from methods other than SuSiE are the same as the results in Fig. 2. For an explanation of the individual plots, see Fig. 2.

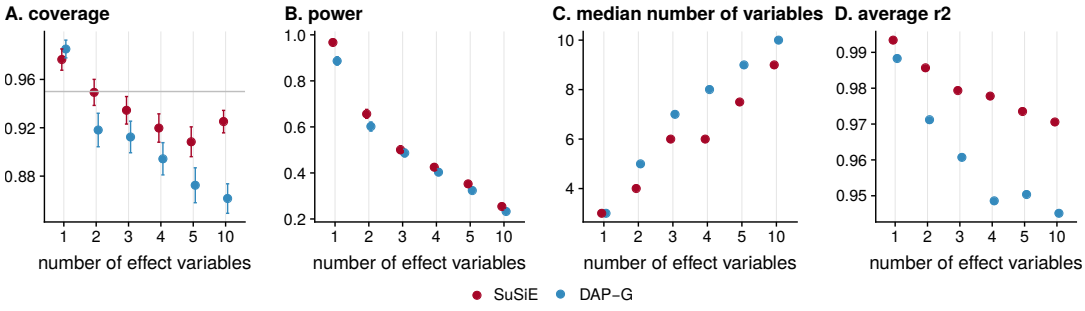

FIGURE S5 Comparison of 95% credible sets (CS) from SuSiE, in which the prior variances σ^2 are estimated rather than fixed to 0.1, and DAP-G: (A) coverage, (B) power, (C) median size, and (D) average squared correlation among variables in each credible set. The DAP-G results shown here are the same as the DAP-G results shown in Fig. 3. For an explanation of the individual plots, see Fig. 3.

REFERENCES

Bertsekas, D. P. (1999) *Nonlinear programming*. Belmont, MA: Athena Scientific, 2nd edn.

Castillo, I., Schmidt-Hieber, J. and van der Vaart, A. (2015) Bayesian linear regression with sparse priors. *Annals of Statistics*, **43**, 1986–2018.

GTEx Consortium (2017) Genetic effects on gene expression across human tissues. *Nature*, **550**, 204–213.

Li, Y. I., van de Geijn, B., Raj, A., Knowles, D. A., Petti, A. A., Golan, D., Gilad, Y. and Pritchard, J. K. (2016) RNA splicing is a primary link between genetic variation and disease. *Science*, **352**, 600–604.

Maller, J. B., McVean, G., Byrnes, J., Vukcevic, D., Palin, K., Su, Z., Howson, J. M. M., Auton, A., Myers, S., Morris, A., Pirinen, M., Brown, M. A., Burton, P. R., Caulfield, M. J., Compston, A., Farrall, M., Hall, A. S., Hattersley, A. T., Hill, A. V. S., Mathew, C. G., Pembrey, M., Satsangi, J., Stratton, M. R., Worthington, J., Craddock, N., Hurles, M., Ouwehand, W., Parkes, M., Rahman, N., Duncanson, A., Todd, J. A., Kwiatkowski, D. P., Samani, N. J., Gough, S. C. L., McCarthy, M. I., Deloukas, P., Donnelly, P. and Donnelly, P. (2012) Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nature Genetics*, **44**, 1294–1301.

Wakefield, J. (2009) Bayes factors for genome-wide association studies: comparison with P-values. *Genetic Epidemiology*, **33**, 79–86.