ESM Methods 3: rjMCMC

We used MCMC methods to explore the space of all potential models (i.e. all possible combinations of features). We therefore used the model-space approach which samples efficiently from trans-dimensional models by allowing for a parameter vector with variable dimension, such that there is no need to sample redundant parameters from vague priors (1–3). Consequently, we only need to define a prior for the regression coefficients $\beta$ and for $k$, the number of variables in our model. We chose a weakly informative normal prior for $\beta$ with a precision of 0.001. A binomial prior was used with the “number of trials” parameter, here representing the total number of features under consideration (3). We set the “success” probability to 0.5 as suggested and motivated e.g. in (4,5). The sampling procedure is then carried out in two separate parts: in the first, discrete part, a new model identifier $M=m'$, corresponding to a set of $k'$ variables is chosen. In the second part, a set of regression coefficients $\beta$ ‘of dimension $k'$ is chosen. For the discrete part, first one of 4 possible moves is chosen: “death”, “birth”, “replace” or “do nothing”. In the case of “death” and “birth” several variables can be added/deleted; these variables are chosen uniformly. The corresponding regression coefficients $\beta$ are then sampled from its full conditional distribution. This analysis was performed using WinBUGS (6) and the R project for statistical computing (http://www.r-project.org/).

In order to assess the robustness of the selected features with respect to the model priors, we repeated the analysis with informative priors favoring sparser models. We therefore set the “success” probability to 0.25 and 0.1 and performed the model selection as described above. We compared resulting feature rankings based on the marginal probabilities for a feature to included in a model as well as the ranking of the best models as these were the main results analyzed in detail in the main paper text. In supplementary figures S1-S3, the robustness of the feature ranking is illustrated: in all cases the ranking is preserved and the core set of 7 features as well as the set of 10 features analyzed in detail are on top of the ranking. In figures S4-S5 it can be seen that a smaller success probability results in sparser models. However, the most important features are consistently selected, and even when favoring very sparse models ($k=0.1$), $HLA$, $PTPN22$, $INS$, $ERBB3$ and $IL2RA$ (part of the core set) are selected in each of the best 20 models.
Supplementary Figure S1: Features ranked according to their marginal probability to be included in the model for a non-informative prior (k=0.5) with all models being equally likely.

Supplementary Figure S2: Features ranked according to their marginal probability to be included in the model for a sparse prior (k=0.25)
Supplementary Figure S3: Features ranked according to their marginal probability to be included in the model for a sparse prior (k=0.1)

Supplementary Figure S4: All accepted models from the rjMCMC algorithm when using a success probability of k=0.25. A colored rectangle is drawn whenever a SNP was included in the respective model. The color code shown at the bottom refers to the log10 odds ratios of the SNP in the model.
Supplementary Figure S5: All accepted models from the rjMCMC algorithm when using a success probability of k=0.1. A colored rectangle is drawn whenever a SNP was included in the respective model. The color code shown at the bottom refers to the log10 odds ratios of the SNP in the model.


