We used a binomial model to calculate the power to discover recurrent events at different population frequencies as a function of patient cohort size and a fixed detection sensitivity of , where is the mis-detection rate, as described in Lawrence et al6. We assume a gene length (the targeted promoter size), a fixed mutation rate of 2.96 mutations/Mb (the average mutation rate) and a value of 1. We then calculate the probability of seeing at least one mutation by chance in each patient as and the signal for each mutation population frequency as When the background mutation rate or mutation frequency is high, this guarantees ; otherwise the equation reduces to (ref. 6).

To determine power, we first calculate the minimal number patients that would reach genome-wide significance, i.e. p-value<0.05/25,000, assuming 25,000 promoters and . The power is then the probability of observing at least patients with a mutation under the alternate model (i.e. a binomial model with ). Smoothed power calculations were performed for constant and variable (Fig. 4a; Extended Data Fig. 9).

Calculation of functional mutation rates (Fig. 4c) was performed assuming total territory for promoters and 75% of the coding gene length for coding genes.