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 $d= 0.44 =\left(1-m\right)$, where $m$ is the mis-detection rate, as described in Lawrence et al6. We assume a gene length $L=650 $(the targeted promoter size), a fixed mutation rate $μ $of 2.96 mutations/Mb (the average mutation rate) and a $f\_{g} $value of 1. We then calculate the probability of seeing at least one mutation by chance in each patient as $p\_{0}=1- \left(1-μf\_{g}\right)^{L} $and the signal for each mutation population frequency $r$ as $p\_{1}=1-\left(1-p\_{0}\right)\*\left(1-r\*d\right). $ When the background mutation rate $μ $or mutation frequency $r$ is high, this guarantees $p\_{1}\leq 1$; otherwise the equation reduces to $p\_{1}≈ p\_{0}+r\*d $(ref. 6).

To determine power, we first calculate the minimal number $n\_{\left\{min\right\}}$ patients that would reach genome-wide significance, i.e. p-value<0.05/25,000, assuming 25,000 promoters and $p=p\_{0}$. The power is then the probability of observing at least $n\_{\left\{min\right\}}$ patients with a mutation under the alternate model (i.e. a binomial model with $p=p\_{1}$). Smoothed power calculations were performed for constant $m$ and variable $r$ (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.