Supporting Methods

Methylation Assay

Genome-wide methylation status was assessed after sham only treatment in four independent matched B-cell/LCL pairs. Strongly unmethylated islands (SUMI) or unmethylated CGs, were assayed as previously described [1, 2]. Briefly, DNA was extracted by standard methods, and digested overnight with the methylation-sensitive restriction enzyme HpaII (New England Biolabs). Indexed sequencing libraries were constructed from the HpaII digest using the standard Illumina kit. For each individual sample, libraries of size 200-350bp and 350-500bp were excised from an agarose gel, PCR-amplified separately and then pooled prior to sequencing. The final insert size range of each pooled library was 80-380bp. Libraries were sequenced on an Illumina GAii to collect paired reads of 50 bases. Only reads passing the Illumina quality filter (chastity filter = 0.6) were further processed. (Table S2) shows the number of reads collected for each sample and passing the quality filter. First we removed fragments that do not contain HpaII restriction ends (CGG) at both ends. Reads were then aligned to the human genome (hg19) using Bowtie [3]. We used an alignment policy that allows up to one mismatch in the first 25 bases and reports only reads that align with a single best match (parameters: -n1 -l25 -m1). Alignments were normalized and analyzed with MetMap to assign to each HpaII site within the scope of the experiment a probability of being unmethylated \( p(U) \) and a probability of being part of an unmethylated region \( p(I) \) [1]. MetMap is available for download at www.cs.berkeley.edu/~meromit/MetMap.html. MetMap annotates strongly unmethylated islands (SUMIs) as regions in which all HpaII sites have a \( p(I) \) greater than 0.1 and at least two HpaII fragments within the region are represented in the MethylSeq data, or by setting a 600 bp interval around each HpaII site that had a \( p(I) \) value smaller than 0.1 and a \( p(U) \) higher than the prior probability of being unmethylated outside of an unmethylated island [4]. We then concatenated all overlapping windows and considered as SUMIs those regions in which at least 30\% of the HpaII sites had a \( p(U) \) larger than the prior-set threshold (0.175), and in which at least two fragments within the region are present in the MethylSeq data. SUMIs share properties with CpG islands [1], but because they are defined by experimental data they are specific to a dataset: the process of annotating SUMIs is most similar to the original definition of CpG islands as HTF (HpaII tiny fragment) islands [5]. While a SUMI is annotated based on the presence of unmethylated HpaII sites in some specimen, it can be scored as methylated if the majority of the HpaII sites within it are methylated. Additionally, since SUMIs are experimentally defined, a region identified as a SUMI in one individual can be methylated in another. SUMIs shared between all 8 samples were cross referenced using BedTools software package [6]. Statistically significant differences in SUMI scores between LCLs and B-cells were identified using paired Student's t-test.

ChIP-seq analysis

EBNA2 and EBNA1 ChIP-seq raw sequence data were reanalyzed from previous studies [7, 8]. Sequences were aligned to the genome using Bowtie2 [9]. Annotation, peak calling and quality control was conducted using the HOMER package [10] with default parameters. Although two EBNA2 datasets were available, one dataset (GSM729851) did not pass quality control, the GC content of the reads was highly skewed relative to dataset 2 and human promoter GC content indicating high level of bias and suggesting compromised experiment and was omitted from the analysis. Only the GSM729852 dataset was used for EBNA2 analysis, however, both available “input” datasets were used as background control.


Figure S1. Hierarchical clustering of all arrays. Hierarchical clustering analysis of the genome-wide expression data that passed all quality controls (15 B-cells, and 11 statin/sham LCLs) demonstrated a distinct clustering by cell type irrespective of treatment status. SCRNA is a control LCL RNA.
Figure S2. Clustering analysis was performed principle components derived from genes that were statin responsive in both B-cells and LCLs (n=173, P<0.05). Principle component analysis was performed using \textit{prcomp} R package, and hierarchical clustering of B-cells and LCLs was performed using the principal components 3-8, which explained 25% of the overall variance. Components 1 and 2 explaining 54% of the variance were excluded as they correlated significantly with EBV treatment (P=2.2E-16) and statin treatment (P=1E-4) covariate respectively.
Figure S3. 

**ChIP-seq.** UCSC Genome Browser visualization of the HNRNPA1 promoter. 

**Boxplot.** HNRNPA1 gene expression. LCL vs. B-cells is significant at $q=3.1\times10^{-14}$. Down-regulation by statin treatment is significant at $q=5\times10^{-4}$.
No overrepresented GO terms

Figure S4. Overlap using genes that are downregulated q <0.05 and EBNA2 peaks, overlap significant P=1*10E-5
Figure S5. (A) The magnitude of change observed in paired B-cells and LCLs is highly correlated in the 173 genes changed with statin treatment in both B-cells and LCLs (unadjusted p<0.05). (B) Genes overlapping between LCL and B-cell statin treatment (p-value <0.05).
Figure S6. Model of MYC/EBNA2/Statin interactions in LCLs. In this model both EBNA2 and MYC influence the expression of alternative splicing factors. Statins inactivate MYC which causes overall down-regulation of alternative splicing factors.
Figure S7. Pairwise scatterplot of simvastatin treated B-cell transcriptome. Outlier S342 was eliminated from downstream the analysis.
Figure S8. Pairwise scatterplot of sham treated B-cells transcriptomes. Outliers S342, S455 and S493B highlighted in yellow were eliminated from downstream analysis.
Figure S9. Pairwise scatterplot of statin treated LCL transcriptomes.
Figure S10. Pairwise scatterplot of sham treated LCL transcriptomes.