ACE: A Probabilistic Model for Characterizing Gene-Level Essentiality in CRISPR Screens

Elizabeth Hutton[1] **and Adam Siepel**[1, ✉]

[1]Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor NY, USA

High-throughput knockout screens using CRISPR-Cas9 are now a widespread method for evaluating the essentiality of genes in different cell types. Here, we introduce a probabilistic model and inference framework, Analysis of CRISPR-based Essentiality (ACE), designed to test for differential signatures of essentiality between cell lines. ACE estimates the essentiality of each gene using a flexible likelihood framework based on the CRISPR-Cas9 experimental process and the observed sequencing counts. In addition, our method can identify which genes are essential only in a specified subset of samples by directly contrasting the likelihood of competing hypotheses — whether a gene has the constant or a differential essentiality between samples. We show using simulations that our approach improves the accuracy of essentiality predictions compared to other methods, and is especially useful for the identification of weaker signals of essentiality. ACE performance was further validated on publicly available CRISPR screen data to distinguish between essential and nonessential genes. In summary, this method provides an improved quantification of essentiality specific to cancer subtypes, and a robust probabilistic framework to identify genes of interest.

CRISPR | essentiality

Correspondence: ***asiepel@cshl.edu***

For decades, high-throughput genetic screens for essential genes were costly and slow in mammalian genomes. With CRISPR-Cas9 technology (Clustered Regularly Interspaced Short Palindromic Repeats — CRISPR-associated protein 9) and large panels of cell lines now available, tissue-specific sensitivities and pathogenic mutations can be tested at scales never before attempted (1–10). Genetic vulnerabilities can now be demonstrated through direct perturbation, rather than relying upon indirect associations through population enrichment studies or expression levels. Phenotypic evaluation across a diverse panel of cell lines allows the early identification of potential tissue-specific side effects, and the isolation of genes essential only to a pathologic biological process. These subtype-specific genetic vulnerabilities are ideal targets for novel therapies guided by personalized genomics.

CRISPR knockout screening offers excellent sensitivity when compared to previous work using RNAi, thanks to CRISPR's effectiveness at disrupting targeted elements and reduced off-target effects (11–13). The CRISPR-Cas9 knockout system in particular allows the design of extremely flexible and high-throughput negative selection assays in human cell lines (Fig. 1). In a typical screen, the artificially expressed Cas9 nuclease introduces a double-stranded break in native DNA, guided by homology with the separately infected single strand guide RNA (sgRNA). The cell then either inserts or deletes a small region of the flanking nucleotides, destroying the target site homology, and repairs the double-stranded break. The resulting genetic lesion disrupts the function of the targeted region. This genetically modified cell proliferates for a series of divisions until harvested by the experimentalist. Any effects on cell growth are measured by sequencing the sgRNA present in the final surviving population of cells. Lethal sgRNA constructs which have introduced a genomic lesion at a site essential for their cell line will be underrepresented in the final population in contrast to nonessential controls. Differential essentiality can be detected from different levels of sgRNA depletion between samples or panels of sample types.

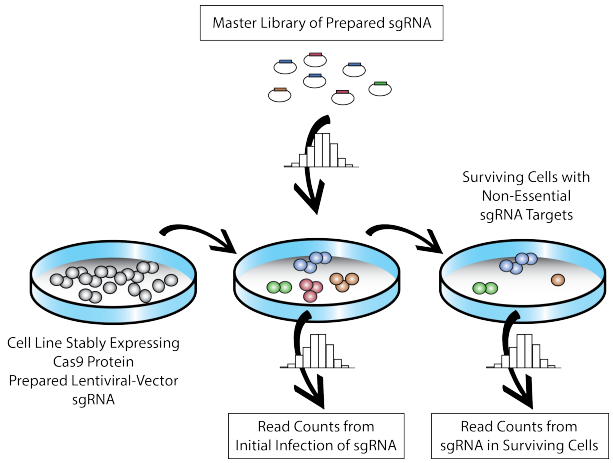


Fig. 1. CRISPR-Cas9 Negative Selection Screen. For each sample, an initial pool of Cas9-expressing cells is infected with a lentiviral sgRNA master library. Depending on the experimental design, either the newly infected population of cells, the original master library, or both are sequenced to measure initial sgRNA abundances. After an experimentalist-determined number of passages, surviving cells are harvested, and the final sgRNA frequencies evaluated.

Many experiments evaluate the loss of cells infected with lethal sgRNA by sequencing both the initial infected and the final population of cells (14–17). Other experiments sequence only the initial pool of sgRNA constructs prior to infection (9, 18, 19), and use the sgRNA abundances in this master library as a proxy for initial frequencies. The master library of sgRNA ranges in abundance across several orders of magnitude, and each biological replicate represents a separate infection of this master library into a new pool of cells.

Most previous methods for statistical inference have relied on summary statistics. However, the use of summary statistics ignores sources of variation unique to each experimental stage, and prohibits adaptation to variation in experimental design (illustrated in Fig. 1). Methods such as BAGEL, JACKS, and others rely upon the log fold change of read abundances (1, 2, 8–

10, 20). Read counts are handled equivalently whether an experiment which has sequenced the initial timepoint or only the master library; if both the master library and initial read counts are available, the user must choose one to use as a 'base'. Other methods simplify gene essentiality into a binary classification of essential or nonessential (2, 4, 21), enabling the identification of clear outliers but neglecting signals of significant but moderate essentiality. Finally, most methods do not test for differential essentiality within sample subtypes (22, 23) (see Fig. S1 for methods comparison).

In this paper, we introduce a model for the Analysis of CRISPR-based Essentiality (ACE), a probabilistic model which allows adaptation to variations in experimental design and principled testing of differential essentiality between sample subtypes. Our approach considers each experimental phase of the CRISPR screen — the master library infection, the initial and final sequencing — as a separate probabilistic process forming a hierarchical Bayesian model. The modularity of our framework permits the analysis of a variety of CRISPR screen databases while accounting for differences in experimental design. Results are tested using both simulations and published sets of essential and nonessential genes. Using a likelihood ratio testing framework, we identify both the degree of a gene's essentiality and the significance of differences between sets of samples.

Results

We have developed a method for the Analysis of CRISPR-based Essentiality (ACE) which leverages information from all available sequenced datasets within a CRIPSR-Cas9 screen, and extracts signal attributable to samples, genes, or individual sgRNA cutting efficiency. By separately modeling each stage of the CRISPR screening process, we can both account for noise unique to this experiment, and address shifting experimental design in this developing field.

ACE Model Overview. ACE exploits the orthogonal sources of information from sample-, gene-, and sgRNA-specific effects by separately inferring parameters for each (see Fig. 2). For every sgRNA g and sample s, users provide sequencing read counts from the master library m_g and/or the initial pool of infected cells x_{sg}, along with the sequenced sgRNA in the depleted pool y_{sg}. Information on each sgRNA's targeted gene and efficiency features must also be provided, along with an estimated number of cells c_s used in the experiment. The model returns sgRNA efficiency, gene essentiality, and the likelihood of differential essentiality between sample subtypes.

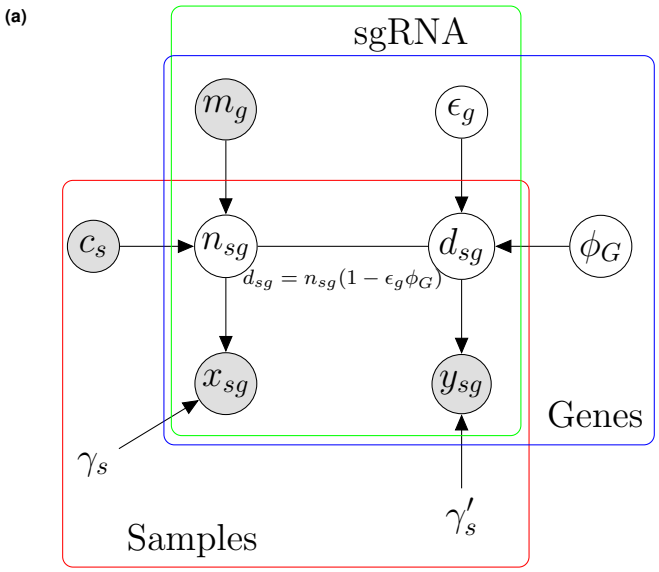
Likelihood Model. . As illustrated in the graphical model shown in Fig. 2, the structure of our likelihood model reflects the experimental design of a CRISPR-Cas9 screen. Each biological replicate s represents an independent sampling of the sgRNA g present in a fraction $m_g \in [0,1]$ of the master library, conditioned upon the number of cells c_s used in the assay. The edited population of cells is then sampled at initial and final timepoints, resulting in the read counts of x_{sg} and y_{sg} respectively. Each of these sampling processes is modeled as a Poisson distribution, resulting in the final likelihood shown in Fig. 2b. As the true abundance of cells in n_{sg} is unknown, we sum over all possible values of $n_{sg} = [0, c_s]$ in the likelihood calculation. The resulting probability density function reflects the known overdispersion of count data (see Fig. S3). With this approach, the ACE model seeks to identify not only the essentiality per targeted element G, but also the experimental parameters unique to CRISPR-Cas9.

Gene-level Essentiality in ACE. We define essentiality ϕ_G as the fractional depletion of cells carrying sgRNAs targeting gene G, after accounting for sgRNA- and sample-specific effects, and ϕ_G ranges from equal to or less than 1 (fully essential). Negative values of ϕ, reflecting increased growth after gene loss, can also be detected (see Methods). The full likelihood of the model pools data across each gene G, each sgRNA g within each gene, and each sample s when inferring model parameters. Gene-level essentiality ($\phi_G \leq 1$, with nonessential genes at 0) is calculated using all samples, and all sgRNAs which target gene G. Each parameter is inferred separately for computational efficiency. The abundance of sgRNA g in the depleted population of cells d_{sg} is evaluated as a function of sgRNA efficiency ϵ_g, initial infected cells n_{sg}, and essentiality ϕ_G (Fig. 2b).

Differential Essentiality Hypothesis Testing. . This probabilistic likelihood framework is used not only to estimate the essentiality of gene G, but also whether this essentiality score ϕ_G is significantly different in different sample types. For each gene G, ACE computes the likelihood that essentiality ϕ_G is constant across all samples. To test for differential essentiality within a sample, separate values of essentiality are estimated for the 'test' sample and for all other samples. The resulting likelihoods compared with the global essentiality likelihood in a likelihood ratio test. This approach directly tests the null hypothesis of whether a uniform essentiality best fits the data, or whether the test and control subtypes rely upon different genes for cell viability and proliferation.

Guide-specific Parameters in ACE. The essentiality of each sgRNA's target locus ϕ_G can only be observed in successfully edited cells. We model cells infected by the sgRNA g as a mixture of effective and ineffective disruptions, with the fraction successfully disrupted described by 'efficiency' $\epsilon_g \in [0,1]$. This efficiency is determined by logistic regression based on a user-defined set of features as in Methods. The efficiency of a class of sgRNA may be predicted by a number of factors, and several groups are actively investigating this area (18, 24–27). Our method estimates efficiency based on sgRNA GC content deciles, though any user-designated categorical sgRNA feature may be used (see Eq. (6)). Other common features used in sgRNA efficiency predictions include proximity to heterochromatin, number of base mismatches, and abundance of off-target sites. No matter the feature, ACE pools information across all genes and samples using logistic regression to identify the relative fraction of successful sgRNA per feature category, ϵ_g, within the context of each experiment.

Sample-specific Scaling Parameters in ACE. The sample-specific scaling factors γ_s and γ'_s affect all sgRNAs and genes



(b)

$$L(\phi_G, \epsilon_g; x_{sg}, y_{sg}, m_g, c_s, \gamma_s, \gamma'_s) = \prod_G \prod_g \prod_s \sum_{n_{sg}=0}^{c_s} P(n_{sg}|c_s m_g) P(x_{sg}|\gamma_s n_{sg}) P(y_{sg}|\gamma'_s n_{sg}(1 - \epsilon_g \phi_G))$$

Fig. 2. Probabilistic Graphical Model of ACE Framework. Each stage of a CRISPR-Cas9 knockout screen is separately modeled with a Poisson sampling process. (a) Graphical model representation of ACE framework. Colored nodes are directly reported in the data; sgRNA efficiency ϵ_g and gene essentiality ϕ_G are inferred, while the number of infected cells n_{sg} is a latent variable integrated out in our likelihood. We assume d_{sg} is defined by the deterministic function $d_{sg} = n_{sg}(1 - \epsilon_g \phi_G)$. γ_s and γ'_s are constant scaling factors calculated prior to parameter inference. (b) Full likelihood equation represented in the graphical model. c_s — Cells infected per sample; d_{sg} — Surviving cells; ϵ_g — Efficiency of sgRNA; ϕ_G — Essentiality of gene G; γ_s, γ'_s — Sample scaling parameters; m_g — Fraction of master library consisting of sgRNA g; n_{sg} — Fraction of cells infected with sgRNA g; x_{sg} — Read counts from initial sequencing; y_{sg} — Read counts from surviving cells.

within a sample. They control for the relationship between the number of cells infected, and the expected sgRNA counts in the initial and final rounds of sequencing, respectively. These scaling parameters are based on a median-of-ratios normalization which reflects the total read count in γ_s, and sample-specific factors in cell growth for γ'_s (see Eq. (4)). For libraries with a large number of essential targets, negative controls are used in place of a median-based normalization. The final pair of sample-specific scaling parameters is independent of the genes targeted by a particular library.

Our essentiality metric ϕ_G pools information across cell lines, and reflects the essentiality common across all samples in a subtype. Thus the panel of samples designated by the user will determine the genetic context of the gene essentiality ϕ_G measured. When choosing a panel of samples, care must be taken to ensure the level of diversity appropriate for the biological question. A varied collection of tissue samples or cell lines is required for the identification of universally essential genes, while a set of patient-derived tissue samples is best for personalized tumor therapeutics. Samples with a genetic background different from the analyzed panel may depend on alternate pathways for cell growth and proliferation, so essentiality values must be generalized with caution.

Simulation Performance. To evaluate the ability of ACE to infer sgRNA-, sample-, and gene-specific parameters, we developed an empirical simulation method to generate read counts from a CRISPR screen with true values precisely defined. To ensure our simulated count data reflects the diversity of observed CRISPR screen data as much as possible, we developed empiriCRISPR, a simple CRISPR screen simulator that models the variation observed in a reference experiment. Our empiriCRISPR simulation framework is independent of parameter inference with ACE, and simulates sgRNA frequencies as a series of three gamma distributions. Final sgRNA frequencies are then scaled and rounded to the user-specified total read counts. Variation and correlation between samples are simulated to match a template CRISPR screen (see Methods). As many biological hyperparameters are difficult to infer, our simulator merely reproduces the variation observed in the final count data from nonessential sgRNA in the template screen (see S 2). All nonessential sgRNA may then be simulated with user-specified effects, providing a gold-standard dataset of known parameters across a range of essentiality values.

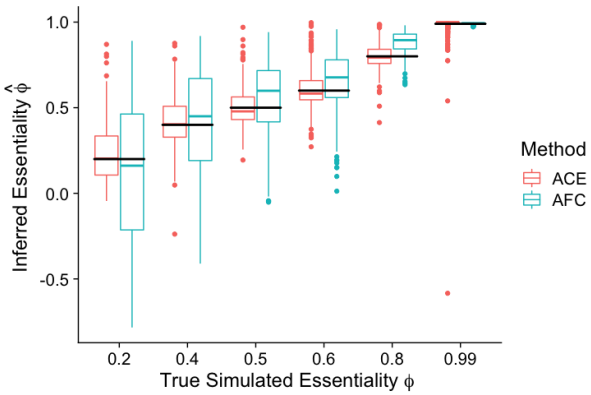


Fig. 3. ACE estimates of simulated essentiality. Counts were simulated for three replicates, with an average of 300 reads per sgRNA. 'ACE' — our probilistic method. 'AFC' — average fold change of sgRNA abundance between initial and final timepoints, and scaled to the essentiality of a set of 100 negative control genes (see Methods). Box plots represent the essentiality estimates from 300 genes simulated at each essentiality value shown; population of 3,150 nonessential genes ($\phi_G = 0$) included in simulation not shown. Black horizontal bars indicate true simulated essentiality.

Simulating Essentiality Gold Standards. To test the capability of ACE for inferring gene essentiality within sample, we simulated a 'gold standard' gene set with empiriCRISPR. Gene essentiality can vary from targets with a complete essentiality $\phi_G = 1$ (100% loss of all associated sgRNA), to nonessential genes with $\phi_G = 0$. To ensure fairness to methods with a median-based normalization method, we simulated 3,150 genes as nonessential ($\phi_G = 0$; 60% of total genes). Three replicates were simulated with an average of 300 reads per sgRNA, with four sgRNA targeting each gene. The noise profile of our simulated counts and relative frequencies of sgRNA in the master library was designed to match the genome-wide CRISPR screens published in Martin et al. 2017 (16), in which both the master library and initial infected sgRNA abundances were sequenced

along with the depleted samples.

Identification of Essentiality Within a Sample. We contrast ACE estimation of essentiality with results derived from the average fold changes as used in studies such as (17, 28) in Fig. 3. ACE infers each gene's essentiality with reduced error, as shown in Fig. S4, and offers an improved estimate of essentiality, with the median estimate closely matching simulated ϕ_G.

Many methods do not provide an estimate of the essentiality of a gene, instead focusing on the binary question of whether or not a gene is essential. We simplified the problem of essentiality inference by simulating 300 genes at 'moderate' ($\phi_G = 0.5$) essentiality, and evaluating the ability of several methods to distinguish essential genes from nonessential controls ($\phi_G = 0$). As shown in Fig. 4, all methods are proficient at identifying essential genes in this simulation.

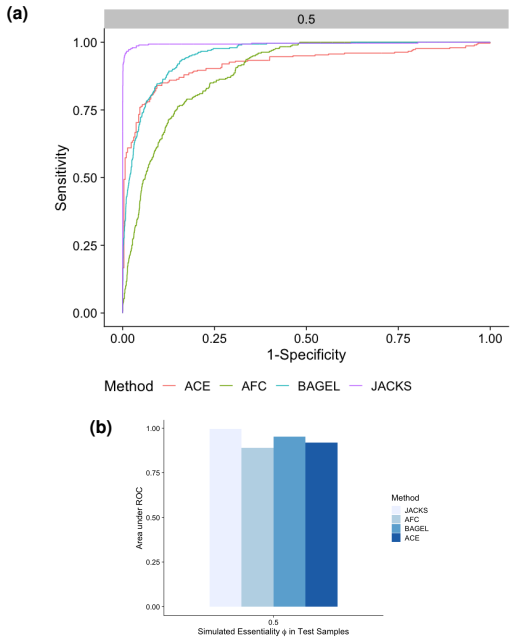


Fig. 4. Binary Essential/Nonessential Classification. While ACE is designed to test for differential essentiality between samples, most methods are designed to test whether a gene is essential or nonessential within a particular sample. Results are shown for the prediction of 'essential' or 'nonessential' using different methods in a sample with 3 replicates and 300 genes simulated $\phi_G = .5$. 3,000 nonessential genes were also simulated, and used as the true positive set in the calculation of the ROC. Only initial and final read counts from the simulation were used for all methods. 'ACE' — likelihood derived from our probabilistic method, ACE. 'AFC' — z-scores calculated from averaging fold changes in sgRNA abundance (see methods). 'BAGEL' — Bayes factor for essentiality reported by the BAGEL method (2). 'JACKS' — p-values for essentiality reported by the JACKS method (8). Fig. 4a ROC curves for all methods at moderate ($\phi_G = 0.5$) and strong ($\phi_G = 0.99$) levels of essentiality. Fig. 4b The area under the ROC curves.

Identification of Differential Essentiality. While measures of gene essentiality ϕ_G within a sample are helpful for identifying sets of essential genes, one advantage of using a statistical method such as ACE lies in the identification of genes with different essentiality between tissue types. These 'differentially essential' genes are vital for exploiting genetic weaknesses in a cancer subtype, or identifying which metabolic pathway a drug resistant cell line is exploiting.

To test ACE performance on the discovery of differential essentiality, we simulated a 'control' and a 'test' sample, each with three replicates. In both samples, uniformly essential control genes were simulated at strong essentiality ($\phi_G = 0.99$, 300 genes), moderate essentiality ($\phi_G = 0.5$, 300 genes), and no essentiality ($\phi_G = 0$, 3,150 genes). In contrast to this 'uniform essentiality' set, genes with differential essentiality were simulated, with essentiality $\phi_{G_{test}} = [.2, .4, .6, .8, .99]$ in the test sample only (see Methods). The ability of ACE and other methods to distinguish differentially essential genes from uniformly essential and nonessential genes is shown in Fig. 5. Using a likelihood ratio test framework, ACE can identify the probability of even moderate differential depletion for each gene, detecting 14% of genes with $\phi_G = .4$ in the test sample only (Fig. S5).

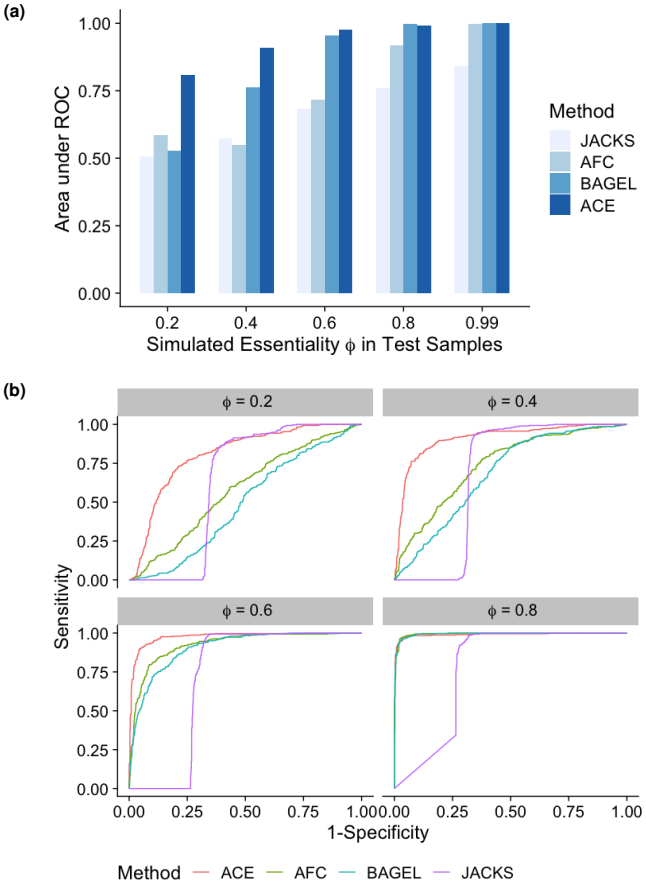


Fig. 5. Comparison of Differential Essentiality Detection. Three replicates of control and test samples were simulated at varying essentiality levels within the test sample ($\phi_{G_{test}} = [.2, .4, .6, .8, .99]$), and $\phi_G = 0$ in the control sample. Uniformly essential controls were simulated at nonessential ($\phi_G = 0$), moderately essential ($\phi_G = 0.5$), and strongly essential ($\phi_G = 0.99$) levels. Fig. 5a Performance identifying 300 genes depleted only in test samples; 300 genes from each of the strong, moderate, and nonessential essentiality control groups were used as true negative controls. Fig. 5b ROC curves for select levels of differential essentiality. Essentiality levels shown are present in the test sample; the same genes are nonessential in the control sample.

Our model simultaneously infers the magnitude of each gene's essentiality within the subset, the significance of that essentiality, and the likelihood that this value of essentiality is unique to the sample subset. Our model may be run with differ-

ent criteria for defining a subset of interest, such as tissue type, tumor-suppressor mutation, or cancer of origin.

ACE Performance on Real Data. We sought to compare the performance of ACE on sets of essential genes identified by previous studies, but a 'gold standard' set has not been established. For our reference essential set, we use 688 genes that have been classified as essential in at least two previously published reference sets as shown in Fig. 6 (1–3). These essential genes are evaluated in published genome-wide screen data, which was also used in the development of this gene set (19). Our nonessential controls are genes with a FPKM < 1 across all cell lines and patient samples in the ENCODE database (see Methods).

When we look at the ability of ACE to re-evaluate essential and nonessential phenotypes within this dataset, we find nonessential genes are consistently fit with lower essentiality values than the essential controls, as expected (see Fig. S6).

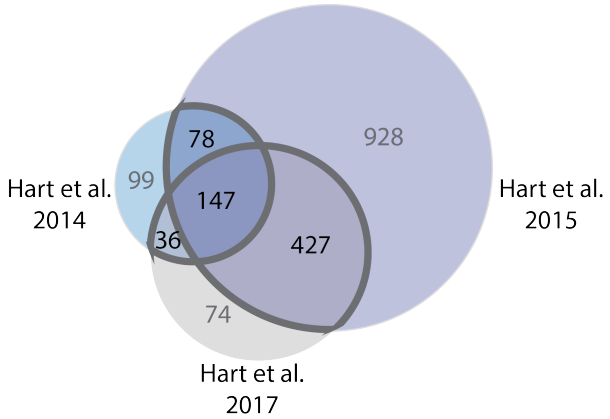


Fig. 6. Selection of Essential Genes. Essential genes were chosen from members of the 'Core Essential Genes' developed by Hart et. al which have been identified as essential in 2 or more publications (2, 3, 29).

Discussion

The rapidly evolving experimental design of CRISPR negative selection screens offers both a challenge and an opportunity for probabilistic modeling. By creating a flexible hierarchical framework, ACE hopes to capture the information specific to CRISPR assays, and provide a rigorous likelihood ratio test for identifying significantly essential genes both within and between samples.

The experimental question driving a CRISPR negative selection screen is not whether or not a gene is essential, but whether the dependencies of one sample type are the same as another's. ACE was designed to isolate these differentially depleted genes within sample subsets, and we see increased sensitivity for finding targets within our simulated data set (see Fig. 5). By combining orthogonal sources of information, ACE is able to identify sgRNA-, sample-, and gene-specific parameters, while providing a principled likelihood framework. This permits the experimental investigator to clearly identify targets of differential essentiality, both in certainty of the essentiality estimate and in degree of difference between test and control samples.

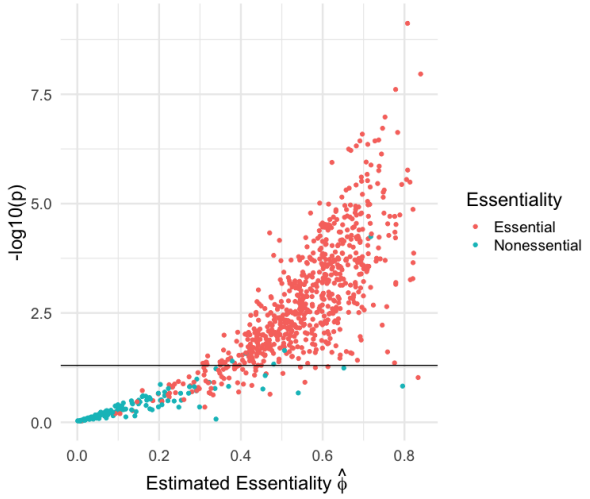


Fig. 7. ACE Performance on Essential Genes. Screening data from the genome-wide CRISPR screen performed by Tzelepis et. al (19) was analyzed using ACE. Shown are essential and nonessential genes (see Methods; p-values fit to ACE likelihood-ratio results using an empirical null distribution based on all nonessential genes). Likelihood of essentiality is based on the likelihood of estimated essentiality $\widehat{\phi_G}$ compared with the likelihood a gene is nonessential ($\widehat{\phi_G} = 0$).

Simulations provide an objective gold standard to evaluate method performance. While sets of core essential genes will continue to be developed as more cell lines are screened with higher quality CRISPR libraries, simulated datasets provide a reliable ground truth. It can be difficult to estimate the false discovery rate of differential essentiality targets as few controls exist, but simulations through empiriCRISPR let us estimate a custom error rate for a particular experiment. In addition to matching CRISPR experimental design, this probabilistic framework provides estimates for essentiality with a robustness to experimental noise that misleads a straightforward averaging of the data (see Fig. 3). As our understanding of CRISPR mechanics continues to evolve, a simulation responsive to shifting sgRNA efficiencies and replicates occludes the necessity of excessive numbers of samples and controls.

As illustrated by the rapid evolution of 'core essential gene' sets over the past 10 years, the definition of 'essential' is a subjective and somewhat arbitrary classification. Essential gene candidates must be tested within each cell line, and then further validated through bespoke knockout experiments. Researchers should be cautious with relying upon published sets of 'essential' gene sets which have not been evaluated within their tissue type of interest, and carefully select the genetic background of the panel of samples used to evaluate essentiality. However, the case of differential essentiality is a relative metric, and will continue to be of interest to researchers exploring genetic vulnerabilities in cancer subtypes and tissue-specific metabolic networks.

Methods

Model Description. The core of our model is a series of Poisson sampling processes integrated over true sgRNA abundance,

with a likelihood ratio test to evaluate significance. Essentiality estimates $\phi_G \in [-\infty, 1]$ are calculated for every genetic element G. Negative values of ϕ_G indicate cells *increase* in abundance after perturbation of the genetic element; positive values indicate the element is essential for cell growth or proliferation. A graphical model representation of ϕ_G is shown in Fig. 2.

While sequenced read counts are usually modeled with a Negative Binomial distribution instead of Poisson (5, 30), we introduce a similar level of overdispersion in our probability density during our integration over unobserved parameters (see Fig. S3)).

The likelihood function used by ACE to derive parameters for gene essentiality ϕ_G and sgRNA efficiency ϵ_g is structured as follows, given master library sgRNA abundances m_g, initial read counts x_{sg}, total cells infected c_s, and depleted read counts y_{sg}:

$$L(\phi_G, \epsilon_g; x_{sg}, y_{sg}, m_g, c_s, \gamma_s, \gamma'_s) = \prod_G \prod_g \prod_s \sum_{n_{sg}=0}^{c_s} P(n_{sg}|c_s m_g) P(x_{sg}|\gamma_s n_{sg}) P(y_{sg}|\gamma'_s n_{sg}(1-\epsilon_g \phi_G)) \quad (1)$$

Where n_{sg} represents true number of cells infected with each sgRNA, and γ_s and γ'_s are sample-specific scaling parameters. Each of the three sections of this likelihood represent an experimental stage in a CRISPR screen. The full likelihood is derived in the following sections.

CRISPR Master Library Normalization. When multiple replicates of the sgRNA master library are sequenced, we normalize each replicate by the total number of reads per sample and average across replicates. The resulting representation of relative master library frequencies, m_g for sgRNA g, is used when modeling the infection of the master library. Alternatively condensed master library read counts may be submitted.

Infection of the sgRNA master library. The prior for the initial abundance n_{sg} of sgRNA g within sample s is described by the following Poisson distribution. The mean μ_1 is determined by the experimentalist-reported total number of cells infected in our screen c_s, and the relative frequency m_g of each sgRNA g in the master library:

$$P(n_{sg}|m_g) = \text{Poisson}(n_{sg}|\mu_1 = c_s m_g) \quad (2)$$

Because the number of infected cells n_{sg} can not be directly observed, our model sums over the range of $n_{sg} \in [0, c_s]$. This summation step is the major bottleneck in our code during parameter optimization, as it must be numerically evaluated for each sgRNA in every sample, and is approximated in intervals of 10 cells.

Initial read counts of sgRNA abundance. The sgRNA sequencing counts x_{sg} from the initial infected population n_{sg} are also modeled as a Poisson sampling process, with scaling parameter γ_s per sample s absorbing the relationship between sequencing depth and number of cells infected:

$$P(x_{sg}|n_{sg}) = \text{Poisson}(x_{sg}|\mu_2 = \gamma_s n_{sg}) \quad (3)$$

Refinement of Sample Scaling Parameters. A change in sgRNA abundance across all sgRNAs within a sample can reflect any number of biological features, from the retardation of cell growth in the presence of Cas9 expression, to variation in passage number between screen designs. Our method addresses these factors, along with variation in sequencing depth, by scaling the number of infected cells n_{sg} with sample parameters γ_s and γ'_s, for the initial and final timepoint reads x_{sg} and y_{sg} respectively.

These sample scaling parameters γ_s and γ'_s are set according to the median ratio of counts per sample to a reference sgRNA abundance. The reference is set either using the frequency of sgRNAs in the master library, or the average abundance of each sgRNA in the initial samples (x_{sg}) if no master library is available. By default only sgRNAs targeting negative control genes (G_{neg}) are used to calculate this ratio, though the median ratio across all sgRNAs may be specified. This metric for read counts is then scaled by c_s, the total number of cells in the experiment. The scaling factor for the final timepoint, γ'_s, is consequently calculated as follows:

$$\gamma'_s = \frac{1}{c_s} \underset{g \in G_{neg}}{\text{median}}\left(\frac{y_{sg}}{m_g}\right) \quad (4)$$

In the case of uniformly essential libraries without negative controls and with identically efficient sgRNA, it is impossible to de-identify sample-specific effects from gene-specific essentiality without the use of positive controls with known values of essentiality.

Modeling sgRNA depletion. In a typical CRISPR negative selection screen, there are two rounds of sequencing. The first set of sequencing counts we model as a Poisson distribution, with a mean of μ_2 as described above. The second round of sequencing takes place after several cell divisions, allowing defects in cell growth and proliferation to manifest. For now, we assume that the abundance of a given sgRNA post depletion, d_{sg}, reflects a deterministic process driven solely by the gene essentiality ϕ_G of the genetic unit G targeted. In all screens examined in this paper, a single coding gene represents a genetic unit. An essentiality value of 0 corresponds with no effect on cell growth or proliferation. Negative essentiality is allowed by our model, observed as an increase in cell growth and proliferation due to perturbation of a deleterious genetic element.

$$P(y_{sg}|n_{sg}(1-\epsilon_g \phi_g)) = \text{Poisson}(y_{sg}|\mu_3 = \gamma'_s n_{sg}(1-\epsilon_g \phi_G)) \quad (5)$$

Guide-specific Effects. This definition of μ_3 considers sample-specific and gene-specific depletion, but CRISPR screens are further affected by the efficiency of the sgRNA chosen. ACE calculates sgRNA-specific efficiency ϵ_g from an independent categorical feature $\bar{F}$. For our simulations, we used the decile of GC content for each sgRNA's template guide. Users may also derive a sgRNA categorical feature through an efficiency score formula such as those developed in (24, 31–35). Guide efficiency ϵ_g is derived from fitting weight ω_f to the feature category F_g of guide g:

$$\epsilon_g \sim \text{logit}(\omega_f F_g) \quad (6)$$

Weight ω_f is estimated by maximum likelihood across all guides within each category in $\bar{F}$. This feature weight allows a fraction of the population of a sgRNA category to be active, but assumes the sgRNA within a particular cell will either be wholly successful or unsuccessful at disrupting its target element. Effective sgRNA drive cell depletion according to gene essentiality ϕ_G, leading to our full description of μ_3 in Eq. (5).

Full Model Likelihood. The full likelihood of the model takes into account all genes G, all sgRNA g within each gene, and all samples s.

$$L(\phi_G, \epsilon_g; x_{sg}, y_{sg}, m_g, c_s, \gamma_s, \gamma'_s) = \prod_G \prod_g \prod_s \sum_0^{c_s} \text{Poiss}\big(n_{sg}|c_s m_g\big)\text{Poiss}\big(x_{sg}|\gamma_s n_{sg}\big) \text{Poiss}\big(y_{sg}|\gamma'_s n_{sg}(1-\epsilon_g\phi_G)\big) \tag{7}$$

As ϕ_G affects the likelihood of the observed counts x_{sg} and y_{sg} across all samples, sample scaling γ'_s influences all genes within each sample, and sgRNA efficiency weights ϵ_g affect classes of sgRNA shared between all genes, concurrent numerical optimization of these three parameters must use data across all genes and samples simultaneously. For tractability with genome-wide screens, we infer ϕ_G and ϵ_g iteratively, permitting parallelization by gene and sgRNA efficiency class. We numerically optimize the following likelihoods to infer gene essentiality ϕ_G and the weight ω_f of each sgRNA efficiency category:

$$\widehat{\phi_G} = \underset{\phi_G}{\text{argmax}}\big\{L(\phi_G; x_{sg}, y_{sg}, m_g, c_s, \gamma_s, \gamma'_s, \epsilon_g = 1)\big\}$$
$$\widehat{\omega_f} = \underset{\omega_f}{\text{argmax}}\big\{L(\epsilon_g = \text{logit}\omega_f F_g; x_{sg}, y_{sg}, m_g, c_s, \gamma_s, \gamma'_s, \widehat{\phi_G})\big\} \tag{8}$$

Weights ω_f are initialized to treat all sgRNA as 100% effective for the first optimization of ϕ_G. To evaluate whether the estimated value of $\widehat{\phi_G}$ represents a significant essentiality, we compare this likelihood to $L(\phi_G = 0)$ using a likelihood-ratio test. P-values for ϕ_G are calculated empirically using a null distribution from the likelihood ratio values from negative control genes.

$$\text{LRT} = \ln \frac{L(\widehat{\phi_G})}{L(\phi_G = 0)} \tag{9}$$

Adaptation to Experimental Design. Some CRISPR screens do not sequence their prepared master library, instead relying on initial and final sequencing to provide information on relative sgRNA abundances. Others do not sequence the initial pool of infected cells, and instead measure only the sgRNA present in the master library. For these experimental modifications, the modularity of our model is quite useful, allowing us to omit the portion of our likelihood for which we have no experimental data.

For experiments without master library information, we use a uniform prior for values of n_{sg}, with our likelihood as follows. Terms μ_2 and μ_3 are given in equations (3) and (5) respectively:

$$L(\phi_G; x_{sg}, y_{sg}) = \prod_G \prod_g \prod_s \sum_{n_{sg}=0}^{c_s} \frac{1}{c_s}\text{Poiss}\big(x_{sg}|\mu_2\big)\text{Poiss}\big(y_{sg}|\mu_3\big) \tag{10}$$

For experiments without sequencing of initial infected sgRNA abundances, we remove the portion of our likelihood addressing x_{sg}:

$$L(\phi_G; m_g, y_{sg}) = \prod_G \prod_g \prod_s \sum_{n_{sg}=0}^{c_s} \text{Poiss}\big(n_{sg}; c_s m_g\big)\text{Poiss}\big(y_{sg}; \mu_3\big) \tag{11}$$

Identification of Subtype-Specific Genetic Vulnerabilities. To identify genes with significant differences in essentiality in one set of samples ($s_t \in S_{test}$) as opposed to a contrasting panel ($s_c \in S_{ctrl}$), we employ a simple likelihood ratio test for every gene G:

$$\begin{aligned} L(\phi_G^{all}) &= L(\phi_G; x_{sg}, y_{sg}, m_g, \epsilon_g, c_s, \gamma_s, \gamma'_s) \\ L(\phi_G^{test}) &= L(\phi_G; x_{s_t g}, y_{s_t g}, m_g, \epsilon_g, c_{s_t}, \gamma_{s_t}, \gamma'_{s_t}) \\ L(\phi_G^{ctrl}) &= L(\phi_G; x_{s_c g}, y_{s_c g}, m_g, \epsilon_g, c_{s_c}, \gamma_{s_c}, \gamma'_{s_c}) \\ \begin{array}{c}\text{LRT for Differential}\\ \text{Depletion of } G\end{array} &= -2\log\frac{L(\phi_G^{all})}{L(\phi_G^{test})L(\phi_G^{ctrl})} \end{aligned} \tag{12}$$

If the test statistic from this likelihood ratio falls significantly outside the empirical null distribution derived from the negative controls, and the associated ϕ_{test} is significantly different from 0 according to Eq. (9), the gene is essential only in the samples within the test subset panel.

All samples are used to calculate sample-specific parameters γ_s and γ'_s, sgRNA-specific parameter ϵ_g, and gene essentiality parameter ϕ_G^{all}. Only ϕ_G is re-estimated for the sample subsets to calculate ϕ_G^{test} and ϕ_G^{ctrl}; estimates for all other parameters are held constant. Our model assumes constant ϵ_g across all samples for all sgRNA with efficiency category F_g.

CRISPR Screen Data and Essential Gene Sets. The essential genes evaluated in Fig. S6 are analyzed using screen data from (19). This experiment screened 5 AML cell lines with 2 replicates per cell line. Only the master library was sequenced, with no sgRNA counts from the initial infection. The genome-wide master library targets 18,010 genes with an average of 5 sgRNA designed per gene. Please see (19) for more experimental details; the CRISPR library is available through Addgene (ID 67989).

The set of essential genes used for validation was derived from the intersection of at least two of the datasets illustrated in Fig. 6 (1–3). The set of nonessential genes was obtained from 604 ENCODE total RNA-seq datasets for 128 cell lines, in vitro differentiated cells, primary cells, and tissues (see Supplemental Table 1). For the nonessential set, we chose genes expressed < 1 FPKM in all cell lines.

Simulation of CRISPR Knockout Screens. As core essential gene sets vary based on experimental design and even between publications (1–3, 9), we developed a series of computational simulations to generate sgRNA read counts under different experimental conditions with known gene essentiality. Each gene was independently simulated using a series of normal distributions to represent each stage of a CRISPR knockout screen (infection, initial and final sequencing). Distribution parameters were tuned to match published CRISPR knockout screens from (16) (see S 2). Initial master library abundances were selected by sampling sgRNA libraries described above. The abundance of each sgRNA construct was assumed to be independent from all other sgRNA used in the experiment. Unless otherwise noted, simulations used 300 genes per essentiality value, 4 sgRNA per gene, and 3 samples.

Package Software. ACE is implemented as an R package (R 3.5.0, (36)) using the *data.table* (37) and *R6* (38) packages. Optimization is performed using R's *optim* package with the 'Brent' method.

Estimation of Essentiality from Mean Log Fold Change. For the 'average fold change' ('AFC') scores, a pseudocount of 0.5 is added to both initial (x_{sg}) and final (y_{sg}) read counts to prevent undefined values. Each sample is then normalized by the total read counts to a total of 10 million reads. After fold changes are calculated, values are adjusted according to the mean fold change in negative controls. Consistent with our ACE notation, essentiality estimates are represented on a continuous scale, where a value of zero has no impact on cell growth or proliferation.

$$\text{AFC} = 1 - \left(\frac{(0.5 + y_{sg}) / \sum_g (0.5 + y_{sg})}{(0.5 + x_{sg}) / \sum_g (0.5 + x_{sg})}\right) \qquad \textbf{(13)}$$

Estimation of Essentiality using BAGEL. Results using BAGEL: Bayesian Analysis of Gene Essentiality (1) were derived using software version 0.91 (last modified 09/2015). Simulation data was analyzed using 4% of genes as positive and negative controls, with simulated essentiality scores of 0.99 (99% depletion) and 0.05 (5% depletion) respectively. Differential essentiality was calculated as the difference of Bayes Factors between separate analysis of test and control sample panels, as performed in (1). For the analysis of real data, we used all genes identified as essential or nonessential in our earlier analysis.

Estimation of Essentiality using JACKS. Results using JACKS: joint analysis of CRISPR/Cas9 knock-out screens (8) calculated p-values on test and control sample panels separately, using the built-in hierarchical prior option and the same negative control genes provided to other methods.

Bibliography

1. Traver Hart and Jason Moffat. BAGEL: a computational framework for identifying essential genes from pooled library screens. *BMC Bioinformatics*, 17(1):164, 2016. ISSN 1471-2105. doi: 10.1186/s12859-016-1015-8.
2. Traver Hart, Megha Chandrashekhar, Michael Aregger, Zachary Steinhart, Kevin R. Brown, Graham MacLeod, Monika Mis, Michal Zimmermann, Amelie Fradet-Turcotte, Song Sun, Patricia Mero, Peter Dirks, Sachdev Sidhu, Frederick P. Roth, Olivia S. Rissland, Daniel Durocher, Stephane Angers, and Jason Moffat. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell*, 163(6):1515–1526, 2015. ISSN 10974172. doi: 10.1016/j.cell.2015.11.015.
3. Traver Hart, Amy Hin Yan Tong, Katie Chan, Jolanda Van Leeuwen, Ashwin Seetharaman, Michael Aregger, Megha Chandrashekhar, Nicole Hustedt, Sahil Seth, Avery Noonan, Andrea Habsid, Olga Sizova, Lyudmila Nedyalkova, Ryan Climie, Leanne Tworzyanski, Keith Lawson, Maria Augusta Sartori, Sabriyeh Alibeh, David Tieu, Sanna Masud, Patricia Mero, Alexander Weiss, Kevin R. Brown, Matej Usaj, Maximilian Billmann, Mahfuzur Rahman, Michael Constanzo, Chad L. Myers, Brenda J. Andrews, Charles Boone, Daniel Durocher, and Jason Moffat. Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3:Genes|Genomes|Genetics*, 7(8):2719–2727, 2017. ISSN 2160-1836. doi: 10.1534/g3.117.041277.
4. Wei Li, Han Xu, Tengfei Xiao, Le Cong, Michael I Love, Feng Zhang, Rafael A Irizarry, Jun S Liu, Myles Brown, and X Shirley Liu. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology*, 15(12):554, dec 2014. ISSN 1474-760X. doi: 10.1186/s13059-014-0554-4.
5. Michael I Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):550, 2014. ISSN 1474-760X. doi: 10.1186/s13059-014-0550-8.
6. Zhiyin Dai, Julie M. Sheridan, Linden J. Gearing, Darcy L. Moore, Shian Su, Sam Wormald, Stephen Wilcox, Liam O'Connor, Ross A. Dickins, Marnie E. Blewitt, and Matthew E. Ritchie. edgeR: a versatile tool for the analysis of shRNA-seq and CRISPR-Cas9 genetic screens. *F1000Research*, 3:95, oct 2014. ISSN 2046-1402. doi: 10.12688/f1000research.3928.2.
7. Jungsik Noh and Beibei Chen. sgRSEA: Enrichment Analysis of CRISPR/Cas9 Knockout Screen Data, 2015.
8. Felicity Allen, Fiona Behan, Anton Khodak, Francesco Iorio, Kosuke Yusa, Mathew Garnett, and Leopold Parts. JACKS: joint analysis of CRISPR/Cas9 knockout screens. *Genome research*, 29(3):464–471, jan 2019. ISSN 1549-5469. doi: 10.1101/gr.238923.118.
9. Robin M. Meyers, Jordan G. Bryan, James M. McFarland, Barbara A. Weir, Ann E. Sizemore, Han Xu, Neekesh V. Dharia, Phillip G. Montgomery, Glenn S. Cowley, Sasha Pantel, Amy Goodale, Yenarae Lee, Levi D. Ali, Guozhi Jiang, Rakela Lubonja, William F. Harrington, Matthew Strickland, Ting Wu, Derek C. Hawes, Victor A. Zhivich, Meghan R. Wyatt, Zohra Kalani, Jaime J. Chang, Michael Okamoto, Kimberly Stegmaier, Todd R. Golub, Jesse S. Boehm, Francisca Vazquez, David E. Root, William C. Hahn, and Aviad Tsherniak. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature Genetics*, 49(12):1779–1784, 2017. ISSN 15461718. doi: 10.1038/ng.3984.
10. Jiyang Yu, Jose Silva, and Andrea Califano. ScreenBEAM: A novel meta-analysis algorithm for functional genomics screens via Bayesian hierarchical modeling. *Bioinformatics*, 32(2):260–267, 2015. ISSN 14602059. doi: 10.1093/bioinformatics/btv556.
11. David W Morgens, Richard M Deans, Amy Li, and Michael C Bassik. Systematic comparison of CRISPR-Cas9 and RNAi screens for essential genes. *Nat Biotechnol*, 34(6):634–636, 2016. doi: 10.1038/nbt.3567.Systematic.
12. Bastiaan Evers, Katarzyna Jastrzebski, Jeroen P.M. Heijmans, Wipawadee Grernrum, Roderick L. Beijersbergen, and Rene Bernards. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nature Biotechnology*, 34(6):631–633, 2016. ISSN 15461696. doi: 10.1038/nbt.3536.
13. Fumitaka Inoue, Martin Kircher, Beth Martin, Gregory M Cooper, Daniela M Witten, Michael T McManus, Nadav Ahituv, and Jay Shendure. A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome research*, 27(1):38–52, jan 2017. ISSN 1549-5469. doi: 10.1101/gr.212092.116.
14. T. Wang, K. Birsoy, N. W. Hughes, K. M. Krupczak, Y. Post, J. J. Wei, E. S. Lander, and D. M. Sabatini. Identification and characterization of essential genes in the human genome. *Science*, 350(6264):1096–1101, 2015. ISSN 0036-8075. doi: 10.1126/science.aac7041.
15. Tim Wang, Haiyan Yu, Nicholas W Hughes, Walter W Chen, Eric S Lander, and David M Sabatini. Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras. *Cell*, 168(5):1–14, 2017. ISSN 0092-8674. doi: 10.1016/j.cell.2017.01.013.
16. Timothy D. Martin, Danielle R. Cook, Mei Yuk Choi, Mamie Z. Li, Kevin M. Haigis, and Stephen J. Elledge. A Role for Mitochondrial Translation in Promotion of Viability in K-Ras Mutant Cells. *Cell Reports*, 20(2):427–438, 2017. ISSN 22111247. doi: 10.1016/j.celrep.2017.06.061.
17. Yusuke Tarumoto, Bin Lu, Tim D.D. Somerville, Yu Han Huang, Joseph P. Milazzo, Xiaoli S. Wu, Olaf Klingbeil, Osama El Demerdash, Junwei Shi, and Christopher R. Vakoc. LKB1, Salt-Inducible Kinases, and MEF2C Are Linked Dependencies in Acute Myeloid Leukemia. *Molecular Cell*, 69(6):1017–1027.e6, 2018. ISSN 10974164. doi: 10.1016/j.molcel.2018.02.011.
18. Andrew J. Aguirre, Robin M. Meyers, Barbara A. Weir, Francisca Vazquez, Cheng Zhong Zhang, Uri Ben-David, April Cook, Gavin Ha, William F. Harrington, Mihir B. Doshi, Maria Kost-Alimova, Stanley Gill, Han Xu, Levi D. Ali, Guozhi Jiang, Sasha Pantel, Yenarae Lee, Amy Goodale, Andrew D. Cherniack, Coyin Oh, Gregory Kryukov, Glenn S. Cowley, Levi A. Garraway, Kimberly Stegmaier, Charles W. Roberts, Todd R. Golub, Matthew Meyerson, David E. Root, Aviad Tsherniak, and William C. Hahn. Genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 targeting. *Cancer Discovery*, 6(8):914–929, 2016. ISSN 21598290. doi: 10.1158/2159-8290.CD-16-0154.
19. Konstantinos Tzelepis, Hiroko Koike-Yusa, Etienne De Braekeleer, Yilong Li, Emmanouil Metzakopian, Oliver M. M. Dovey, Annalisa Mupo, Vera Grinkevich, Meng Li, Milena Mazan, Malgorzata Gozdecka, Shuhei Ohnishi, Jonathan Cooper, Miten Patel, Thomas McKerrell, Bin Chen, Ana Filipa Filipa Domingues, Paolo Gallipoli, Sarah Teichmann, Hannes Ponstingl, Ultan McDermott, Julio Saez-Rodriguez, Brian J.P. J.P. Huntly, Francesco Iorio, Cristina Pina, George S. S. Vassiliou, Kosuke Yusa, Etienne De Braekeleer, Yilong Li, Emmanouil Metzakopian, Oliver M. M. Dovey, Annalisa Mupo, Vera Grinkevich, Meng Li, Milena Mazan, Malgorzata Gozdecka, Shuhei Ohnishi, Jonathan Cooper, Miten Patel, Thomas McKerrell, Bin Chen, Ana Filipa Filipa Domingues, Paolo Gallipoli, Sarah Teichmann, Hannes Ponstingl, Ultan McDermott, Julio Saez-Rodriguez, Brian J.P. J.P. Huntly, Francesco Iorio, Cristina Pina, George S. S. Vassiliou,

and Kosuke Yusa. A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Reports*, 17(4):1193–1205, 2016. ISSN 22111247. doi: 10.1016/j.celrep.2016.09.079.

20. Gaoxiang Jia, Xinlei Wang, and Guanghua Xiao. A permutation-based non-parametric analysis of CRISPR screen data. *BMC Genomics*, 18(1):545, dec 2017. ISSN 1471-2164. doi: 10.1186/s12864-017-3938-5.
21. Philipp N. Spahn, Tyler Bath, Ryan J. Weiss, Jihoon Kim, Jeffrey D. Esko, Nathan E. Lewis, and Olivier Harismendy. PinAPL-Py: A comprehensive web-application for the analysis of CRISPR/Cas9 screens. *Scientific Reports*, 7(1), dec 2017. ISSN 20452322. doi: 10.1038/s41598-017-16193-9.
22. Hyun-Hwan Jeong, Seon Young Kim, Maxime W C Rousseaux, Huda Y Zoghbi, and Zhandong Liu. Beta-binomial modeling of CRISPR pooled screen data identifies target genes with greater sensitivity and fewer false negatives. *Genome research*, 29(6):999–1008, jun 2019. ISSN 1549-5469. doi: 10.1101/gr.245571.118.
23. Wei Li, Johannes Köster, Han Xu, Chen-hao Chen, Tengfei Xiao, Jun S Liu, Myles Brown, and X Shirley Liu. Quality control , modeling , and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biology*, pages 1–13, 2015. ISSN 1474-760X. doi: 10.1186/s13059-015-0843-6.
24. John G Doench, Nicolo Fusi, Meagan Sullender, Mudra Hegde, Emma W Vaimberg, Katherine F Donovan, Ian Smith, Zuzana Tothova, Craig Wilen, Robert Orchard, Herbert W Virgin, Jennifer Listgarten, and David E Root. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, 34(2):184–191, 2016. ISSN 1087-0156. doi: 10.1038/nbt.3437.
25. Luca Pinello, Matthew C. Canver, Megan D. Hoban, Stuart H. Orkin, Donald B. Kohn, Daniel E. Bauer, and Guo Cheng Yuan. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nature Biotechnology*, 34(7):695–697, 2016. ISSN 15461696. doi: 10.1038/nbt.3583.
26. Martin Jinek, Fuguo Jiang, David W Taylor, Samuel H Sternberg, Emine Kaya, Enbo Ma, Carolin Anders, Michael Hauer, Kaihong Zhou, Steven Lin, Matias Kaplan, Anthony T Iavarone, Emmanuelle Charpentier, Eva Nogales, and Jennifer a Doudna. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science (New York, N.Y.)*, 343(6176): 1247997, mar 2014. ISSN 1095-9203. doi: 10.1126/science.1247997.
27. Robert L. Unckless, Andrew G. Clark, and Philipp W. Messer. Evolution of Resistance Against CRISPR/Cas9 Gene Drive. *Genetics*, 205(2):827–841, feb 2017. ISSN 0016-6731. doi: 10.1534/GENETICS.116.197285.
28. Edwin H Yau, Indrasena Reddy Kummetha, Gianluigi Lichinchi, Rachel Tang, Yunlin Zhang, and Tariq M Rana. Genome-Wide CRISPR Screen for Essential Cell Growth Mediators in Mutant KRAS Colorectal Cancers. *Cancer research*, 77(22):6330–6339, sep 2017. ISSN 1538-7445. doi: 10.1158/0008-5472.CAN-17-2043.
29. Traver Hart, Kevin R Brown, Fabrice Sircoulomb, Robert Rottapel, and Jason Moffat. Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Molecular systems biology*, 10(7):733, 2014. ISSN 1744-4292. doi: 10.15252/msb.20145216.
30. Chen-hao Chen, Wei Li, Tengfei Xiao, Han Xu, Peng Jiang, Clifford A Meyer, and Myles Brown. Integrative analysis and refined design of CRISPR knockout screens. *bioRxiv*, pages 617–632, 2017. doi: http://dx.doi.org/10.1101/106534.
31. Robin Graf, Xun Li, Van Trung Chu, and Klaus Rajewsky. sgRNA Sequence Motifs Blocking Efficient CRISPR/Cas9-Mediated Gene Editing. *Cell Reports*, 26(5):1098–1103.e3, jan 2019. ISSN 2211-1247. doi: 10.1016/J.CELREP.2019.01.024.
32. Santiago Gisler, Joana P. Goncalves, Waseem Akhtar, Johann de Jong, Alexey V. Pindyurin, Lodewyk F. A. Wessels, and Maarten van Lohuizen. Multiplexed Cas9 targeting reveals genomic location effects and gRNA-based staggered breaks influencing mutation efficiency. *Nature Communications*, 10(1):1598, dec 2019. ISSN 2041-1723. doi: 10.1038/s41467-019-09551-w.
33. Sebald AN Verkuijl and Marianne G Rots. The influence of eukaryotic chromatin state on CRISPR–Cas9 editing efficiencies. *Current Opinion in Biotechnology*, 55:68–73, feb 2019. ISSN 0958-1669. doi: 10.1016/J.COPBIO.2018.07.005.
34. Jennifer Listgarten, Michael Weinstein, Benjamin P. Kleinstiver, Alexander A. Sousa, J. Keith Joung, Jake Crawford, Kevin Gao, Luong Hoang, Melih Elibol, John G. Doench, and Nicolo Fusi. Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nature Biomedical Engineering*, 2(1):38–47, jan 2018. ISSN 2157-846X. doi: 10.1038/s41551-017-0178-6.
35. Belen Gutierrez, Jérôme Wong Ng, Lun Cui, Christophe Becavin, and David Bikard. Genome-wide CRISPR-Cas9 screen in E. coli identifies design rules for efficient targeting. *bioRxiv*, page 308148, apr 2018. doi: 10.1101/308148.
36. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.
37. Matt Dowle and Arun Srinivasan. *data.table: Extension of 'data.frame'*, 2019.
38. Winston Chang. *R6: Classes with Reference Semantics*, 2017.

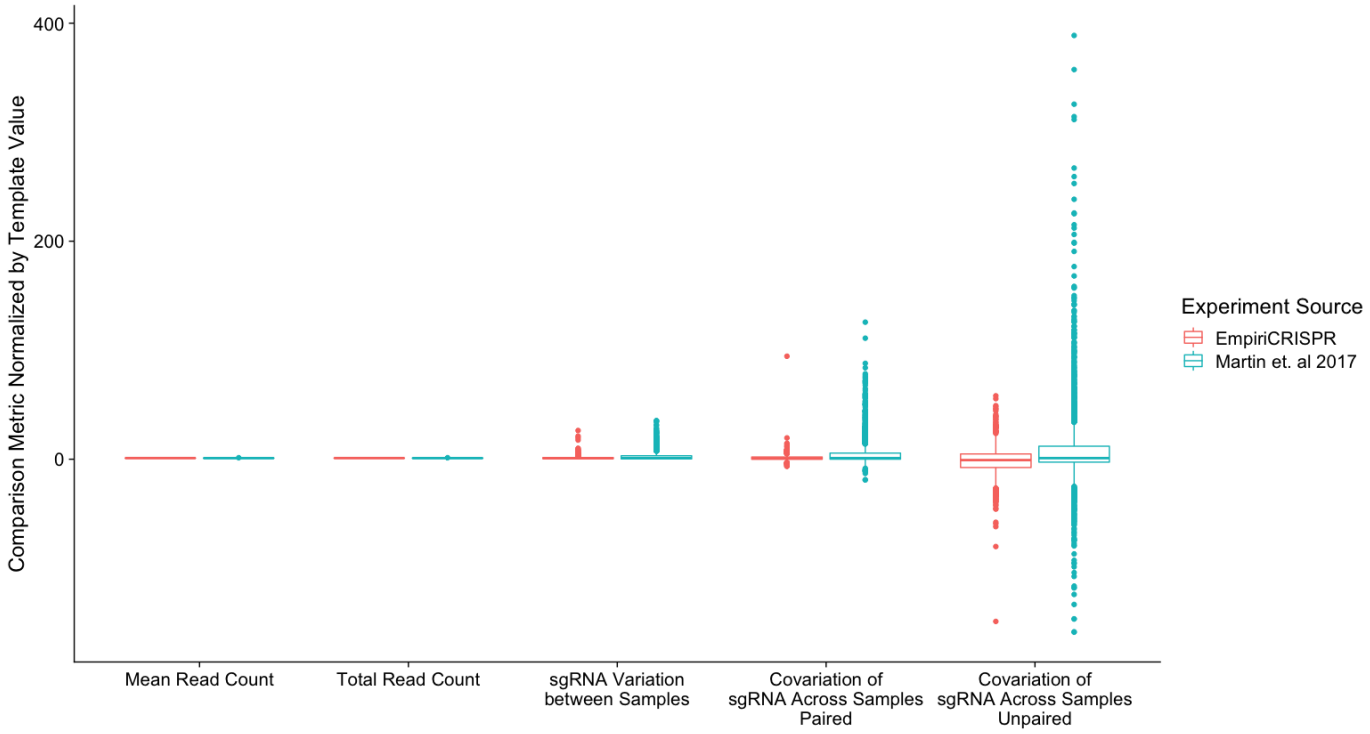
Supplementary Section 1: Motivation for ACE Framework

Method	Directly Models Count Data	Estimates Essentiality	Tests for Differences Between Subtypes
ACE	✓	✓	✓
Average Fold Change (AFC)	✗	✓	✗
BAGEL (Hart et al. 2015)	✗	✗	✓
JACKS (Allen et al. 2019)	✗	✓	✗
MAGeCK-RRA (Li et al. 2014)	✗	✗	✗
CERES (Meyers et al. 2017)	✗	✓	✓
ScreenBEAM (Yu et al. 2015)	✗	✓	✗
CB2 (Jeong et al. 2019)	✓	✗	✗
PinAPL-Py (Spahn et al. 2017)	✓	✗	✗
MAGeCK-MLE (Li et al. 2015)	✓ *	✓	✗

*depleted read counts only

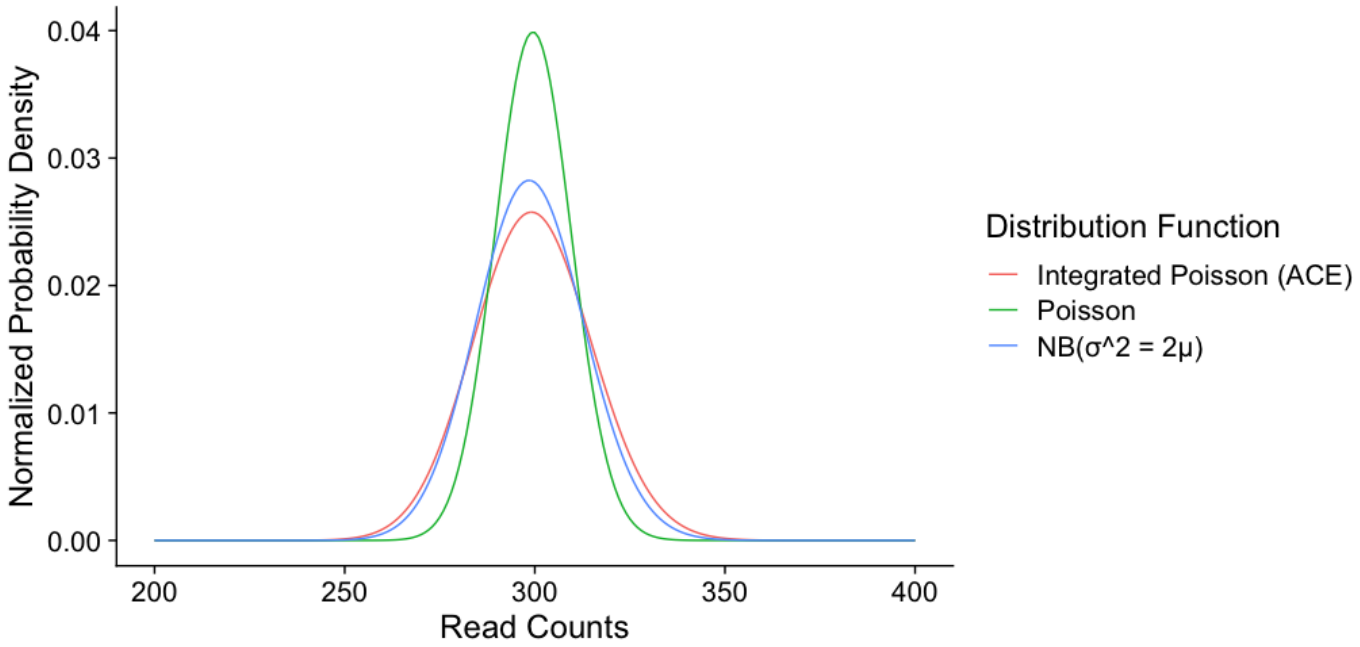
Supplementary Figure S1. Methods for Identifying Differential Essentiality in CRISPR Screens. Many methods have been developed to analyze CRISPR screening data, but none are designed to test for differential essentiality between sample subtypes as a part of their inference framework. Some, such as BAGEL and CERES, test for significant differences in essentiality after parameter inference, but do not test whether the data is best described by a single essentiality parameter ('Tests for Differences Between Subtypes' column). Directly modeling count data, as opposed to relying upon summary statistics such as log fold change, enables ACE to adapt to variations in experimental design, and leverage information from all available data ('Directly Models Count Data' column). Providing a quantification of essentiality, as opposed to fitting a binary essential/nonessential classification, enables researchers to more intuitively estimate phenotype ('Estimates Essentiality' column).

Supplementary Section 2: Description of Simulated Data



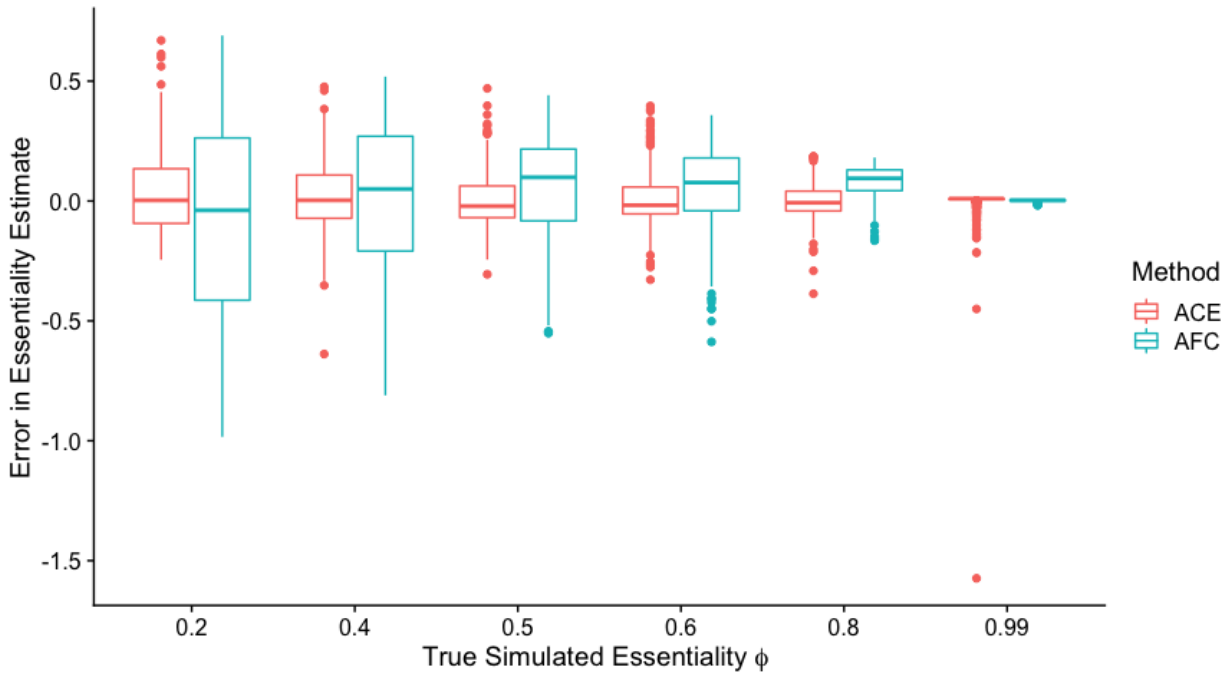
Supplementary Figure S2. Simulated Data for Evaluation of ACE Performance. The genome-wide CRISPR screen performed by Martin et. al 2017 (16) was used as our EmpiriCRISPR template, as this work sequenced the sgRNA master library used as well as the initial and final samples. The sequenced master library provided the initial sgRNA abundances in the simulation, and all metrics shown were calculated in the initial and depleted counts. Shown are approximately 2300 guides targeting globally nonessential genes, as identified using ENCODE expression data (see Methods), and a randomly selected matching subset of nonessential sgRNA from simulated read counts. Mean and total read counts were calculated within each sample, while variation and covariation metrics were calculated for each guide across samples. 'Paired' indicates covariance was calculated between initial and depleted samples, with each sgRNA paired with the depleted sample produced by the same master library transfection; 'Unpaired' have been paired with samples from a different transfection replicate.

Supplementary Section 3: Illustration of ACE Distribution Function



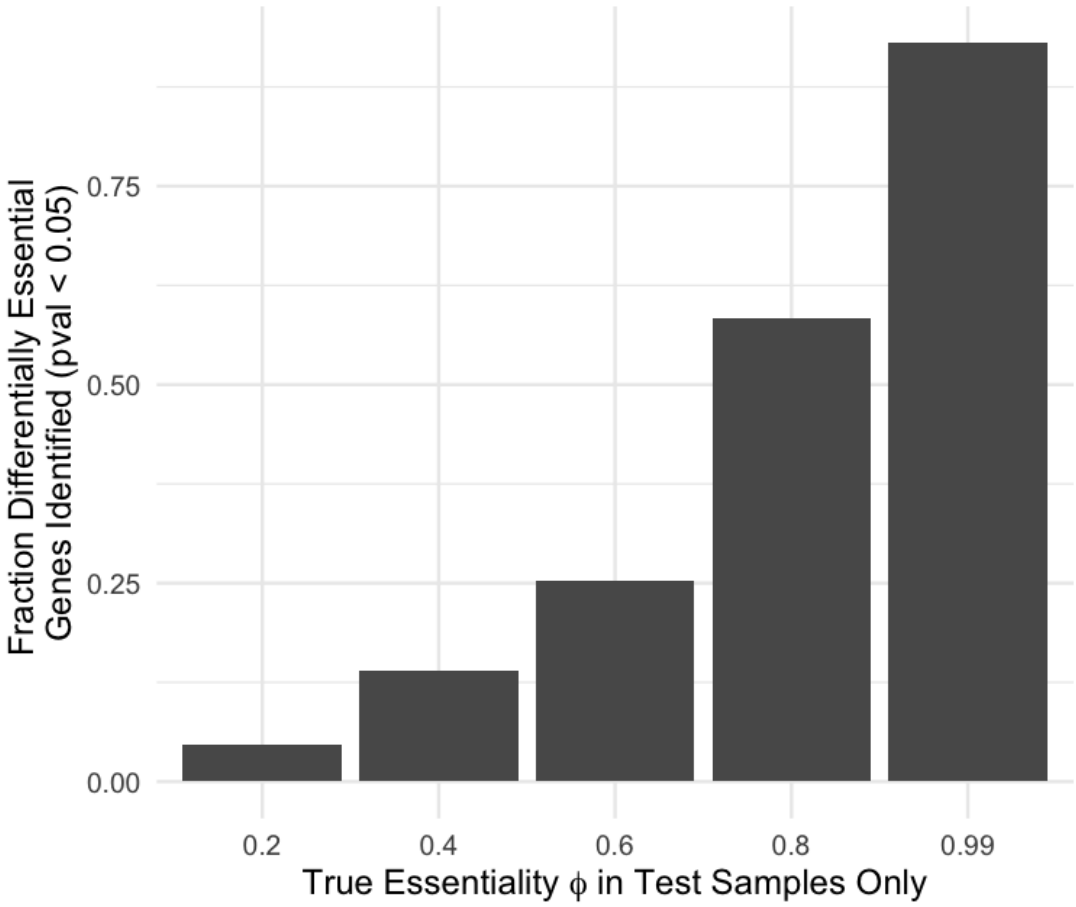
Supplementary Figure S3. Illustration of Dispersion of Different Probability Distributions. Read count data is known to have a high dispersion similar to a Negative Binomial distribution (5). The values shown here reflect the probability density of different values of read counts in a CRISPR screen, as described by a Poisson distribution with point estimates of the mean (green), a Negative Binomial distribution with point estimates of the mean and a variance equal to twice the mean (blue). The ACE model integrates over all possible numbers of infect sgRNA, resulting in a probability density (red) similar to the Negative Binomial.

Supplementary Section 4: Error in Estimation of Essentiality



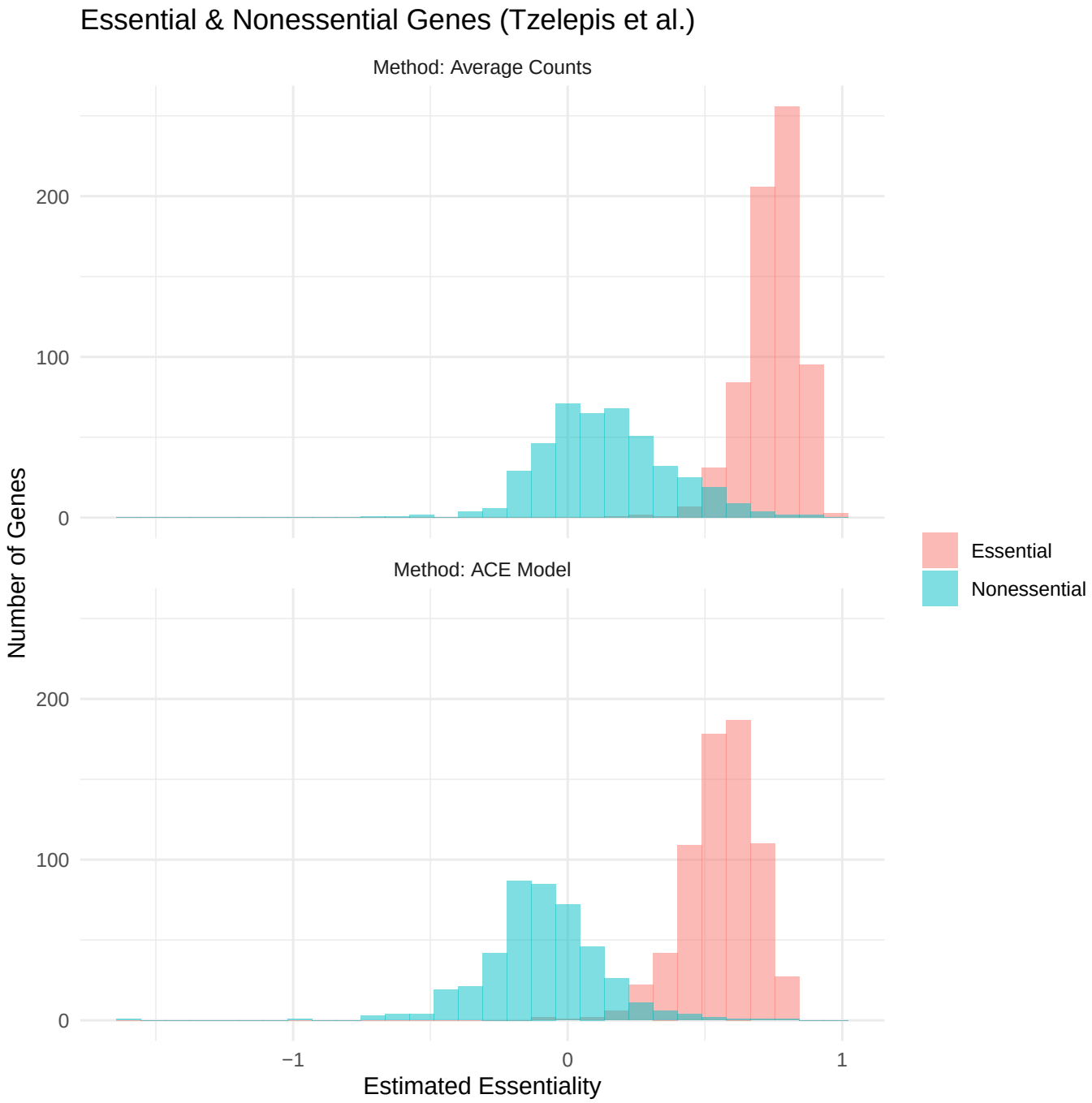
Supplementary Figure S4. Error in Estimation of Gene Essentiality. Gene essentiality ϕ_G was calculated using ACE and average fold change (AFC; see Methods) across three replicates of simulated count data. 300 genes were simulated at each essentiality level shown; results from an additional 3,150 nonessential genes are not shown.

Supplementary Section 5: Fraction of Differentially Essential Genes Detected by ACE



Supplementary Figure S5. Differential Essentiality Prediction. ACE identification of genes essential specific to a user-defined subset of samples. 300 genes were simulated at each essentiality level within a 'test' panel of 3 replicates, while the same genes were simulated as nonessential in 3 control replicates. Shown are the fraction of total genes that are essential only in test samples that were detected as significantly differentially essential ($p < 0.05$) by ACE.

Supplementary Section 6: ACE Performance on Strongest Signals of Essentiality



Supplementary Figure S6. Analysis of Essential Genes from (19). Histogram of essentiality values $\widehat{\phi_G}$ inferred by ACE and averaging sgRNA fold changes (see Methods) for well-supported essential and nonessential genes, as identified by previous literature and ENCODE expression data respectively.