Supplemental Information

Genome-Wide Specificity of DNA-Binding, Gene Regulation, and Chromatin Remodeling by TALE- and CRISPR/Cas9-Based Transcriptional Activators

Lauren R. Polstein[1], Pablo Perez-Pinera[1,#], D. Dewran Kocak[1], Christopher M. Vockley[2,3], Peggy Bledsoe[2], Lingyun Song[2], Alexias Safi[2], Gregory E. Crawford[2,4,*], Timothy E. Reddy[2,5,*], and Charles A. Gersbach[1,2,6,*]

[1]Department of Biomedical Engineering, Duke University, Durham, North Carolina, United States of America, 27708
[2]Center for Genomic and Computational Biology, Duke University, Durham, North Carolina, United States of America, 27708
[3]Department of Cell Biology, Duke University Medical Center, Durham, North Carolina, United States of America, 27710
[4]Department of Pediatrics, Division of Medical Genetics, Duke University Medical Center, Durham, North Carolina, United States of America, 27710
[5]Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina, United States of America, 27710
[6]Department of Orthopaedic Surgery, Duke University Medical Center, Durham, North Carolina, United States of America, 27710

[#]Current address: Department of Bioengineering, 1270 Digital Computer Lab, University of Illinois at Urbana-Champaign, 1304 W. Springfield Ave. Urbana, Illinois 61801 USA

[*]Co-corresponding authors

Address for correspondence:

Gregory E. Crawford, Ph.D.
Center for Genomic and Computational Biology
Duke University
Durham, NC 27708
919-684-8196
greg.crawford@duke.edu

Timothy E. Reddy, Ph.D.
Center for Genomic and Computational Biology
Duke University
Durham, NC 27708
919-684-3286
tim.reddy@duke.edu

Charles A. Gersbach, Ph.D.
Department of Biomedical Engineering
Room 136 Hudson Hall,
Box 90281
Duke University
Durham, NC 27708-0281
919-613-2147
charles.gersbach@duke.edu

Table of Contents:

Supplemental Table 9. DNA-binding sites of dCas9-VP64 targeted to the *IL1RN* gene as determined by ChIP-Seq.

Supplemental Table 10. DNA-binding sites of TALE-VP64s targeted to the *HBG1/2* genes as determined by ChIP-Seq.

Supplemental Table 11. DNA-binding sites of dCas9-VP64s targeted to the *HBG1/2* genes as determined by ChIP-Seq.

Supplemental Table 12. Closest transcriptional start sites (TSS) of genes neighboring off-target ChIP-seq sites for TALE-VP64s targeted to the *IL1RN* gene.

Supplemental Table 13. Closest transcriptional start sites (TSS) of genes neighboring off-target ChIP-seq sites for gRNAs targeted to the *IL1RN* gene.

Supplemental Table 14. Closest transcriptional start sites (TSS) of genes neighboring off-target ChIP-seq sites for TALE-VP64s targeted to the *HBG1/2* gene.

Supplemental Table 15. Closest transcriptional start sites (TSS) of genes neighboring off-target ChIP-seq sites for gRNAs targeted to the *HBG1/2* gene.

Supplementary Table 16. Top 100 differential Dnase hypersensitive sites in cells treated with TALEs targeted to *IL1RN* compared to control.

Supplementary Table 17. Top 100 differential Dnase hypersensitive sites in cells treated with TALE-VP64s targeted to *IL1RN* compared to control.

Supplementary Table 18. Top 100 differential Dnase hypersensitive sites in cells treated with TALEs targeted to *HBG1/2* compared to control.

Supplementary Table 19. Top 100 differential Dnase hypersensitive sites in cells treated with TALE-VP64s targeted to *HGB1/2* compared to control.

Supplementary Table 20. Top 100 differential Dnase hypersensitive sites in cells treated with dCas9 targeted to *IL1RN* compared to control.

Supplementary Table 21. Top 100 differential Dnase hypersensitive sites in cells treated with dCas9-VP64 targeted to *IL1RN* compared to control.

Supplementary Table 22. Top 100 differential Dnase hypersensitive sites in cells treated with dCas9 targeted to *HBG1/2* compared to control.

Supplementary Table 23. Top 100 differential Dnase hypersensitive sites in cells treated with dCas9-VP64 targeted to *HBG1/2* compared to control.

Supplemental Table 24. Closest transcriptional start sites (TSS) of genes neighboring the top 100 DNase-seq sites for TALE-VP64s targeted to the *IL1RN* gene.

Supplemental Table 25. Closest transcriptional start sites (TSS) of genes neighboring the top 100 DNase-seq sites for gRNAs targeted to the *IL1RN* gene.

Supplemental Table 26. Closest transcriptional start sites (TSS) of genes neighboring the top 100 DNase-seq sites for TALE-VP64s targeted to the *HBG1/2* gene.

Supplemental Table 27. Closest transcriptional start sites (TSS) of genes neighboring the top 100 DNase-seq sites for gRNAs targeted to the *HBG1/2* gene.

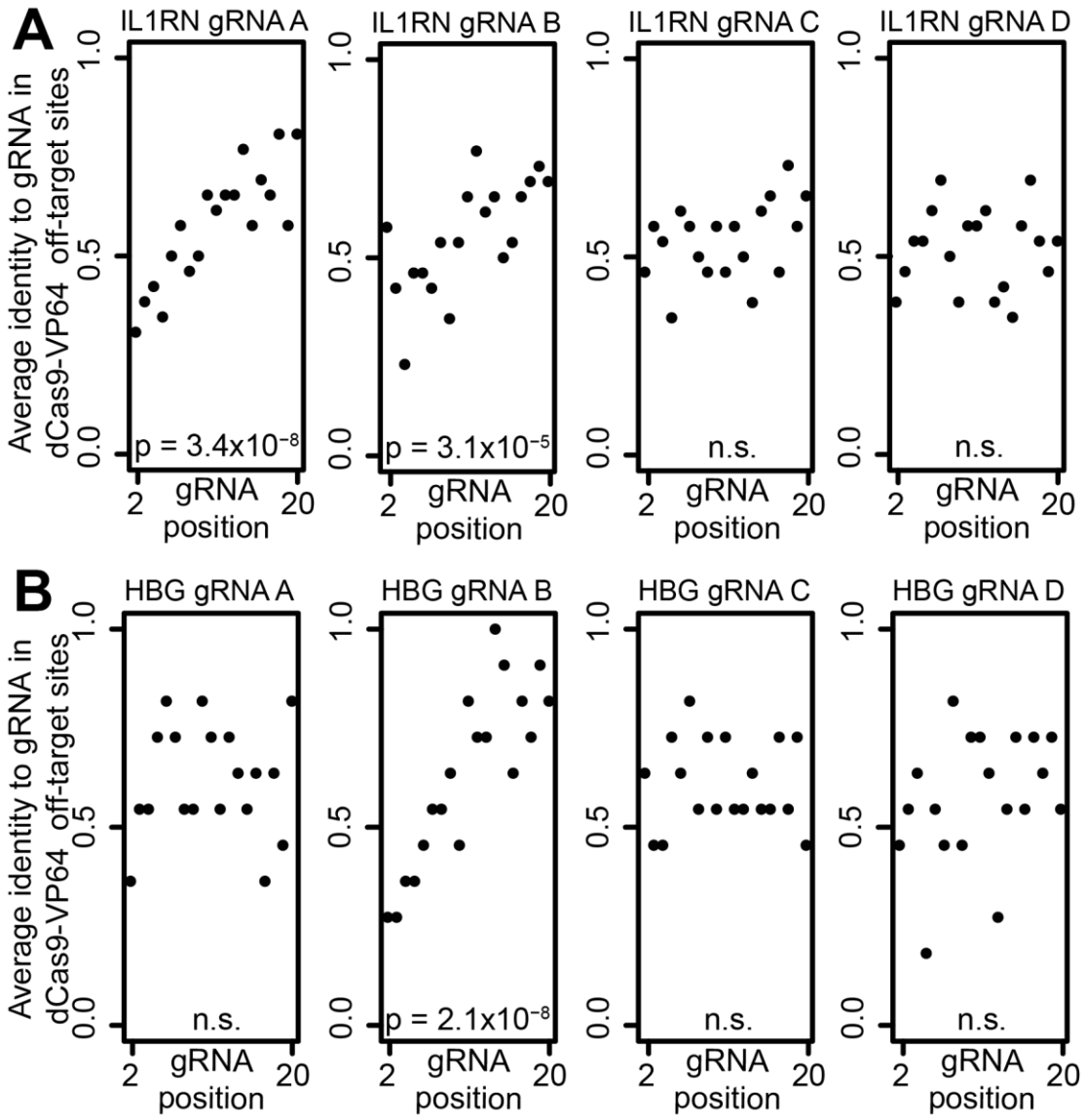
Supplemental Table 28. DNase-seq data at the ChIP-seq off-target sites identified in samples treated with TALE-VP64 targeted to *IL1RN* (Supplementary Table 8).

Supplemental Table 29. DNase-seq data at the ChIP-seq off-target sites identified in samples treated with dCas9-VP64 targeted to *IL1RN* (Supplementary Table 9).

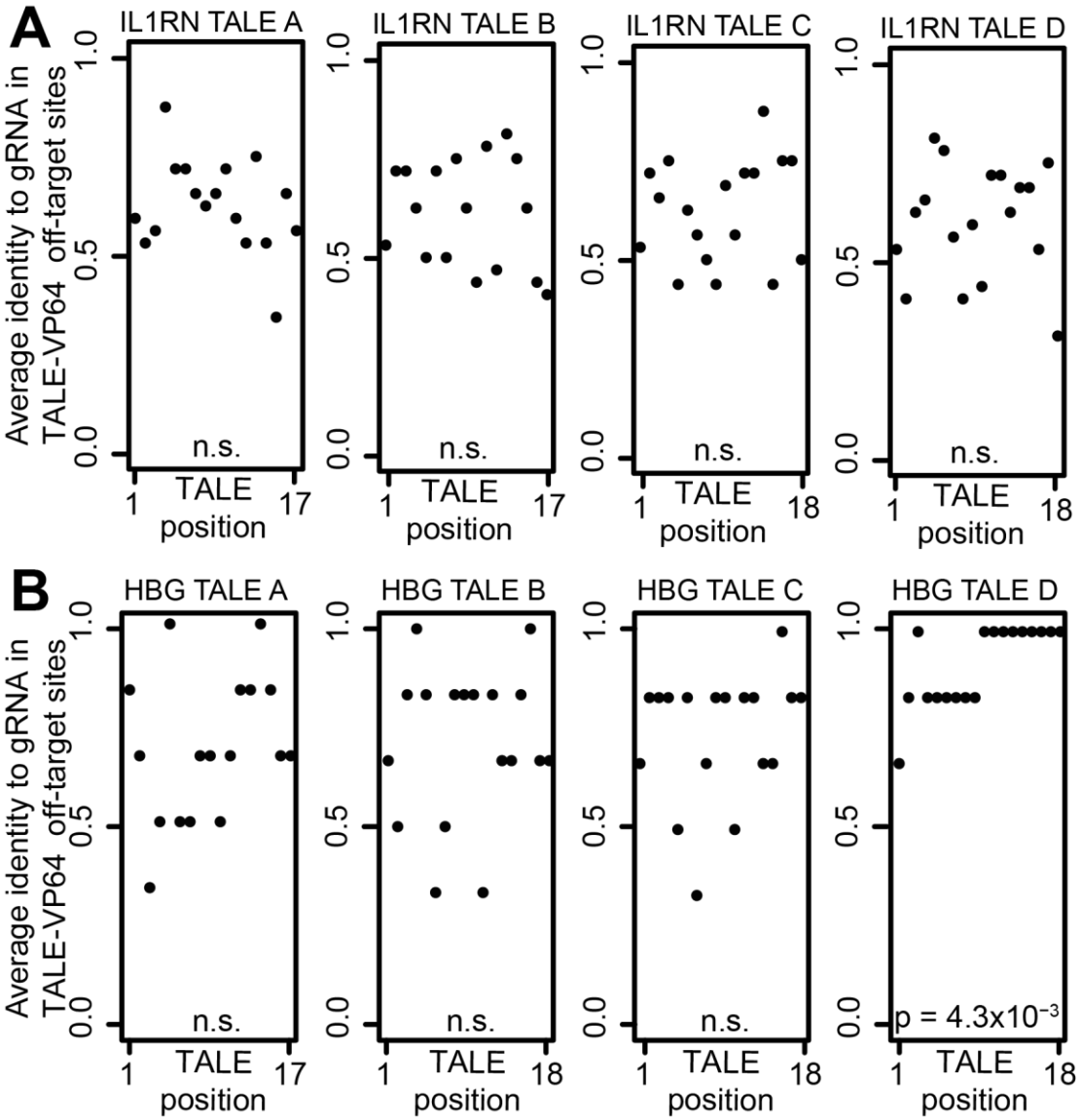
Supplemental Table 30. DNase-seq data at the ChIP-seq off-target sites identified in samples treated with TALE-VP64 targeted to *HBG1/2* (Supplementary Table 10).

Supplemental Table 31. DNase-seq data at the ChIP-seq off-target sites identified in samples treated with dCas9-VP64 targeted to *HBG1/2* (Supplementary Table 11).

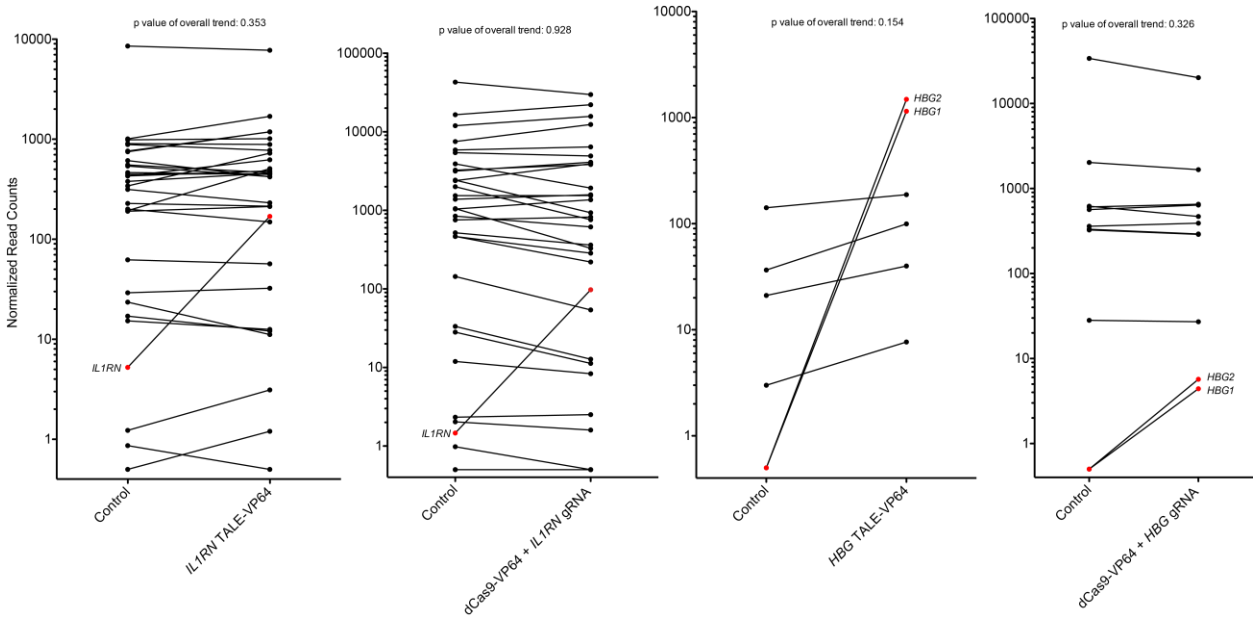
Supplemental Table 32. Number of sequencing reads for DNase-seq analysis for all samples.



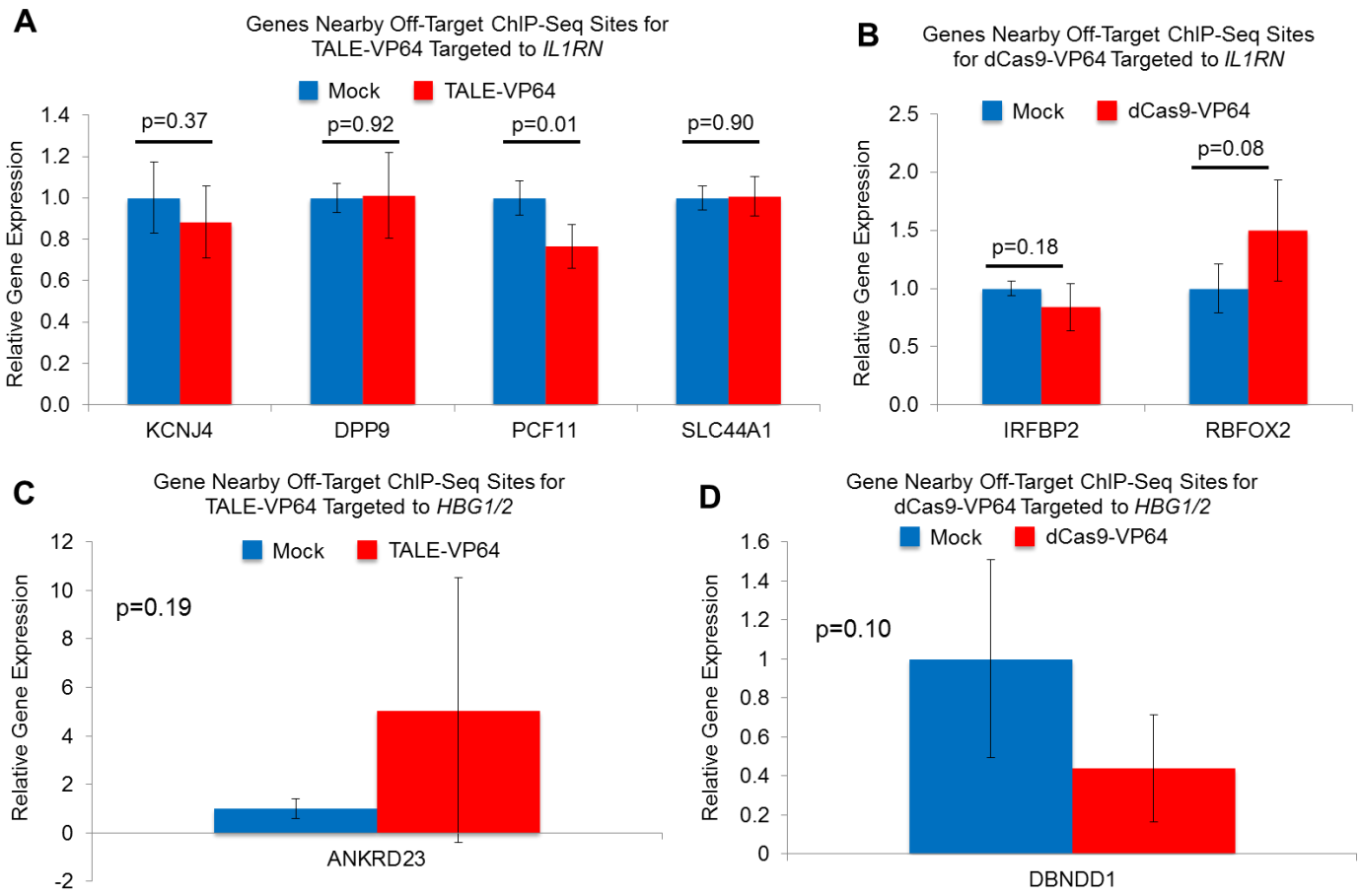
Supplemental Figure 1. Positional trends in similarity for all gRNA target sequences with off-target binding sites. Plots for *IL1RN* gRNA A, *IL1RN* gRNA B, and *HBG* gRNA B are the same as in Figure 4. The additional panels are gRNAs for which an enriched motif was not detected in *de novo* searches. No significant trend was observed for the gRNAs without an enriched motif, further indicating that those gRNAs did not substantially contribute to off-target binding.



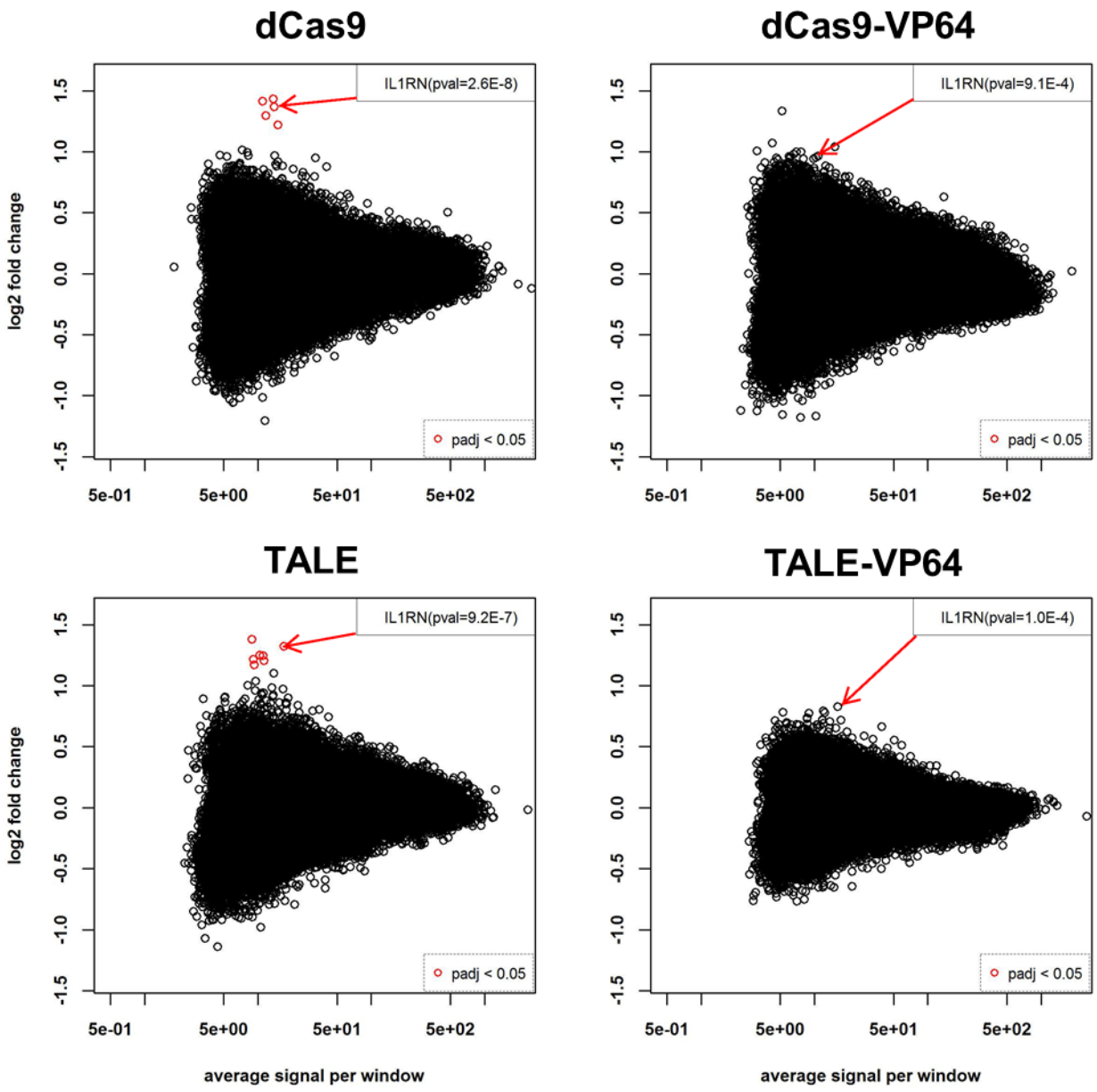
Supplemental Figure 2. Positional trends in similarity for all TALE target sequence in off-target binding sites. The plot for *HBG* TALE D is the same as in Figure 4. The additional panels are TALEs for which an enriched motif was not detected in *de novo* searches, none of which had a significant 3′ trend in sequence similarity to the best match in off-target binding sites.



Supplemental Figure 3. Analysis of expression levels of genes nearby ChIP-seq off-target sites, as determined by RNA-seq. The nearest refseq-annotated transcription start site was determined for each ChIP-seq off-target site using the bedtools software (**Supplemental Tables 12-15**). The read counts for each refseq mRNA, after normalizing for the total number of aligned reads per experiments, were then collected and plotted using Prizm (Graphpad). To determine if there was a significant trend in gene expression across all candidate off-target genes, an ANOVA analysis was used that took into account per-gene mean expression values. No significant trends were observed. All statistical analysis was performed in R.

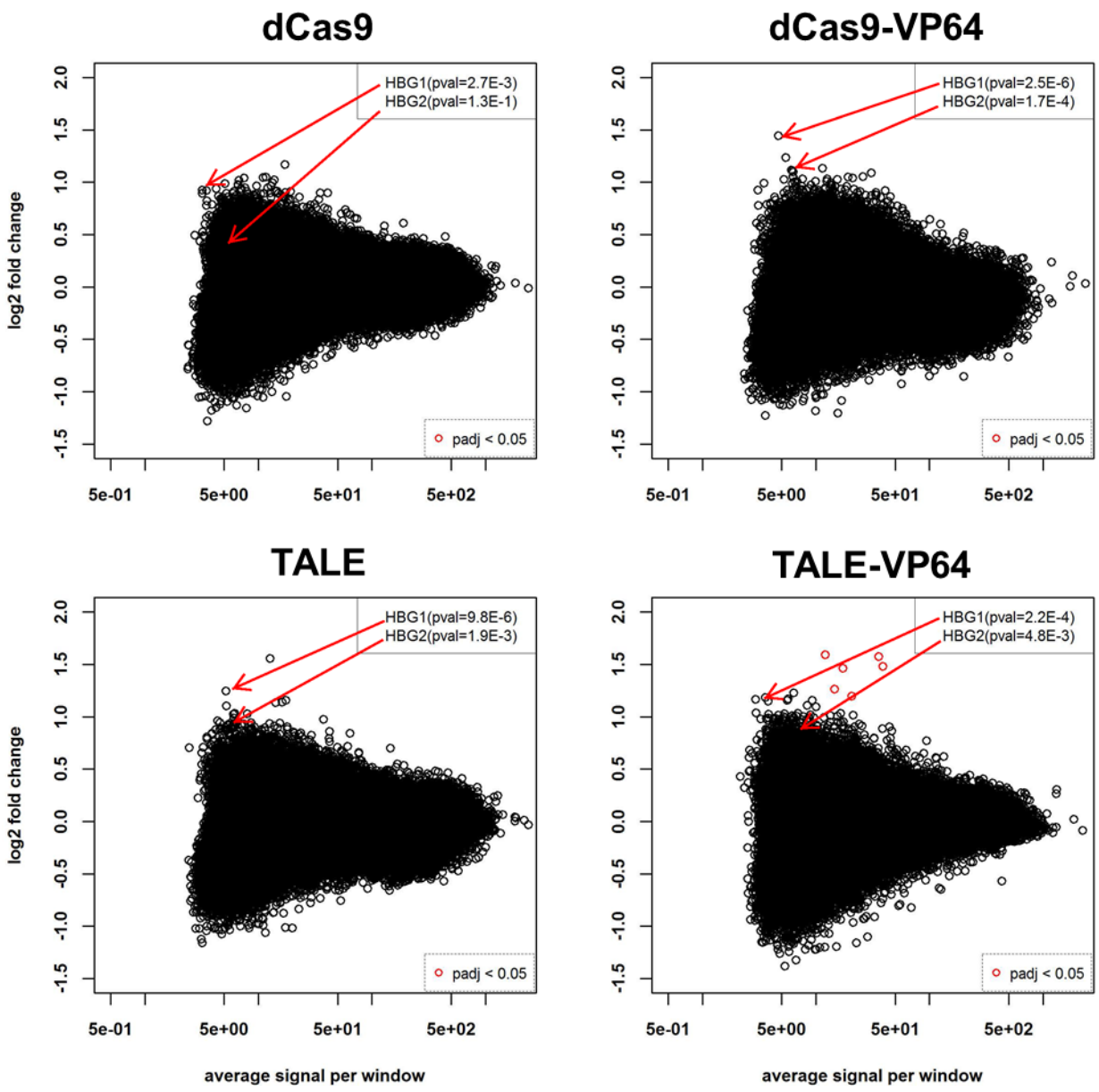


Supplemental Figure 4. The expression of eight representative genes identified nearby ChIP-seq off-target sites was assessed by qRT-PCR in samples treated with the indicated TALE-VP64 or dCas9-VP64/gRNA constructs (n=4, mean ± st. dev.). Relative gene expression levels are calculated by the ΔΔCt method normalized to mock-transfected cells and GAPDH. No significant increases in gene expression were observed.



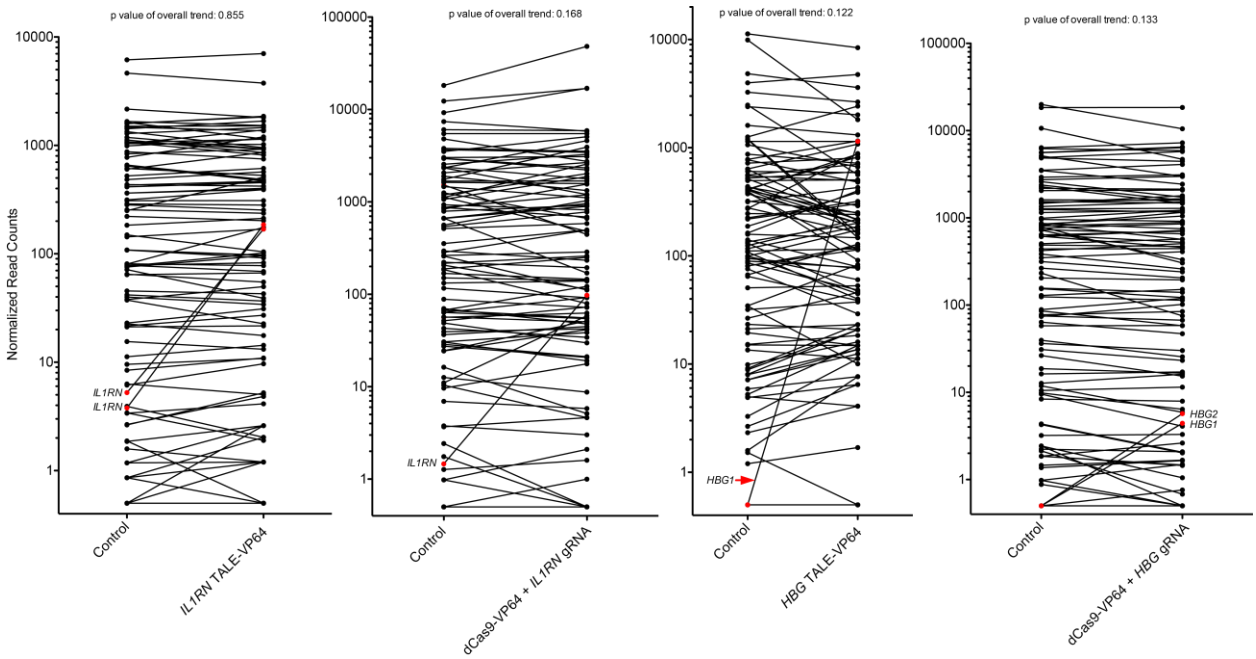
Supplemental Figure 5. DNase-seq results for each individual treatment targeted to the *IL1RN* promoter compared to control. Each dot represents an individual DNase I hypersensitive site. The x-axis is the average signal and the y-axis represents the log fold change compared to mock-transfected control cells as determined by DESeq (Anders and Huber, 2010). Nominal *P* values (pval), and *P* values adjusted for multiple hypotheses testing (False Discovery Rate (padj), shown in red) indicate that *IL1RN* targeting is highly specific. The top 100 differential DNase I hypersensitive sites are listed in **Supplemental Tables 16-17 and 20-21**.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome biology *11*, R106.



Supplemental Figure 6. DNase-seq results for each individual treatment targeted to the *HBG1/2* promoter compared to control. Each dot represents an individual DNase I hypersensitive site. The x-axis is the average signal and the y-axis represents the log fold change compared to mock-transfected control cells as determined by DESeq (Anders and Huber, 2010). Nominal *P* values (pval), and *P* values adjusted for multiple hypotheses testing (False Discovery Rate (padj), shown in red) indicate that *HBG1/2* targeting is highly specific. The top 100 differential DNase I hypersensitive sites are listed in **Supplemental Tables 18-19 and 22-23**.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome biology *11*, R106.



Supplemental Figure 7. Analysis of expression levels of genes nearby DNase-seq off-target sites, as determined by RNA-seq. The nearest refseq-annotated transcription start site was determined for each of the top 100 DNase-seq off-target site using the bedtools software (**Supplemental Tables 24-27**). The read counts for each refseq mRNA, after normalizing for the total number of aligned reads per experiments, were then collected and plotted using Prizm (Graphpad). The on-target genes are labeled in red. To determine if there was a significant trend in gene expression across all candidate off-target genes, an ANOVA analysis was used that took into account per-gene mean expression values. No significant trends were observed. All statistical analysis was performed in R.