Direct profiling of genome-wide dCas9 and Cas9 specificity using ssDNA mapping (CasKAS). Response to reviewer comments

Reviewer 1:

> Marinov et al. describe CasKAS, a method for defining off-target binding and cleavage by mapping of unwound ssDNA regions of cellular genomes. The experimental protocol workflow is attractive because it is relatively quick to perform. The method is a novel off-target activity detection method that detects open R-loops associated with CRISPR-Cas editors by chemically labelling of guanine bases present in unwound ssDNA with N3-kethoxal, based on recently described KAS-seq (kethoxal-assisted ssDNA sequencing). However, some limitations may arise from biological noise caused by endogenous cellular transcription and widespread ssDNA content in cells. The authors test CasKas with four sgRNAs, two mouse and two human. The paper is clearly written, but there are some gaps in the experimental design and presentation. Some weaknesses of the manuscript include: no validation of off-target activity by targeted sequencing for indel mutations in edited versus unedited cells, superficial comparisons to earlier methods, and characterization of a limited number of gRNAs. Overall, this is a solid manuscript that presents limited preliminary characterization of a novel method for detecting the genome-wide unwinding activity of Cas9 in vitro and in human cells. The manuscript could be strengthened by addressing the major and minor questions and concerns I list below.

We thank the reviewer for the comments and suggestions.

> 1. What percentage of CasKAS detected sites can be validated in cells? Are the 198 new sites detected for mouse sgRNA #1 bona fide sites of cellular off-target nuclease activity? An experimental approach to validating these sites (such as by using multiplex targeted sequencing approaches such as rhAMP-seq or others) would be an important first step towards understanding the performance characteristics of CasKas. Targeted sequencing validation should be performed at 6 or more sgRNAs to have a generalizable sense of CasKAS performance and validation rate.

We appreciate the reviewer's concern and have carried out the suggested experiments and included the results in the revised manuscript.

However, we would like to clarify a few points.

Edits can only be validated using sequencing for Cas9 and after editing *in vivo* in live cells.

CasKAS can be used for that purpose, but that is only a fraction of its applications.

CasKAS can be used to map dCas9 occupancy sites inside cells. As this leaves no edits, there is nothing to validate by sequencing.

More importantly, as our results also indicate, the off-target sites identified *in vivo* are often a small subset of the biochemically possible off-target sites. In our opinion, it is better for that more expansive set to be known and considered, as just because they do not show up at a high enough frequency in sequencing validation datasets after *in vivo* editing, it does not necessarily follow that they are never actually occupied and cleaved under any conditions inside cells.

When mapping off-target sites *in vitro* with dCas9, all of the sites that show the proper strand asymmetry around a peak expected from true occupancy events do appear to contain a cognate sequence for the sgRNA, as our original analsyis clearly showed, i.e. it is very hard to see them as spurious off-targets.

When mapping off-target sites *in vitro* with Cas9, we have two kinds of events – occupancy without cutting (where, again, we see cognate sgRNA sequence matches), and occupancy with cleavage, which is clearly evident in the sequence profiles and using our C-score metric. If cutting is observed, then presumably editing will happen too. Thus observing cleavage (and not just cleavage, but cleavage precisely at the basepair in the sgRNA sequence match that it is expected to occur) in the *in vitro* active CasKAS data is already more than sufficient evidence that the site is a real target on its own. This was the reason for our omission of a targeted sequencing experiment in the original version of the manuscript.

2. How does CasKAS generally compare against earlier methods such as GUIDE-seq, Digenome-seq, DISCOVER-seq, and CHANGE-seq? In particular, what proportion of validated sites detected by other methods can also be detected by CasKAS? The comparisons to earlier methods should be performed more broadly with larger numbers of gRNAs.

We have carried out the requested analysis using external datasets where possible.

We used the EMX1 and VEGFA sgRNAs as these are the ones that have been used in most other publications. It is unfortunately not possible for us to include more sgRNAs in such a comparison as we do not have public data for all different assays for a larger number of sgRNAs, e.g. the Digenome-seq paper only provides data for the VEGF-A sgRNA, the CHANGE-seq (which we already compared against) only provides data for the EMX1 sgRNA, etc.

3. Are the comparisons with other methods performed appropriately? For example, while it is reasonable to use ChIP-seq peak calling when making the ChIP-seq comparisons, comparisons with other methods such as CHANGE-seq (Supp. Fig. 16) should use the off-target calls from the original authors (or generated by the associated software) rather than applying ChIP-seq peak calling to non-ChIP-seq data. When making the comparison in Supp. Fig. 16 it would be better to compare against bona fide validated off-target activity rather than predicted off-targets.

The reason we used MACS2 to identify CHANGE-seq enrichment sites is that neither in the Supplementary Information of the CHANGE-seq paper, nor in the SRA accession where we found the EMX1 dataset could we find a list of off-target sites provided by the authors themselves. MACS2 is widely used not only for ChIP-seq but also for other enrichment assays, such as ATAC-seq, so it is unlikely that its use resulted in a dramatic misidentification of enriched sites.

We have now compared our results to those using other assays based on the published lists of off-target sites where possible.

4. Why map the ssDNA-containing regions rather than directly pull down the bound DNA? The rationale for this should be more clearly articulated in the introduction of the manuscript. There are some clear advantages for doing so (potentially more relevant to Cas9 cleavage activity than simply binding) but also other limitations (background of exposed ssDNA due to DNA replication/transcription, etc.). Data that directly compares binding versus CasKAS unwound signal at multiple sgRNA targets, such as could be generated using a biotinylated dCas9 (or alternately Cas9 dChIP-seq, albeit potentially with more noise) would be particularly insightful.

We have edited the manuscript to make these points more clear.

To reiterate, the rationale behind CasKAS is that it is a versatile, simple and quick assay that can map the occupancy of Cas9, dCas9 and all other CRISPR proteins and their derivatives in a wide variety of contexts (*in vitro* and *in vivo*) and using relatively few cells when doing so *in vivo*.

In practical terms this means that CasKAS is much cheaper (an antibody that is enough for at best 20 pull downs costs ~$500, while for the same cost one can buy enough kethoxal for ~400 CasKAS experiments), much faster (the crosslink reversal portion of ChIP-seq alone takes longer than the whole CasKAS protocol), and with many fewer cells (our *in vivo* CasKAS was generated from 400,000 cells; similarly successful dCas9 ChIP-seq would require ≥20M cells)

Scientifically, the ssDNA capture provides information about productive engagement of the Cas9 enzyme with DNA, which is more specific than general physical association (as demonstrated by our comparison against existing ChIP-seq data), and also allows intermediates to be captured, such as the immediate post-cleavage state, all with the same experimental approach.

5. How scalable is CasKAS? For a method which is pitched as being simple and robust, only a small number of sgRNAs have been evaluated. The workflow seems like it should be compatible with higher-throughput characterization and it would be good to see CasKAS applied to at least 10-15 sgRNAs.

We have included in the revised manuscript the application of CasKAS to more sgRNAs as requested.

We note, however, that the goal of the manuscript was to be a proof-of-principle demonstration of the various applications of CasKAS (for which the experiments that we carried out were more than sufficient), not to derive generalizable insights about CRISPR occupancy (for which 10-15 sgRNAs would not be sufficient anyway).

6. Can Cas-Kas be performed without manual curation of peaks and is the software to perform Cas-Kas site detection available? For scientists to use the CasKAS method it will be important to have some objective criteria (implementable in a software pipeline) to define peaks. Otherwise, the sets of peaks would likely vary widely between individual investigators. The source code for the analytical pipeline for the data presented in the paper should be provided for rigor and reproducibility.

We used MACS2, a standard tool for calling regions of enrichment in *-seq assays. No new code for peak calling was written. The scripts for making tracks and other postprocessing were already linked in the manuscript (`https://github.com/georgimarinov/GeorgiScripts`), and are previously described in detail in several publications (PMID: 33606259 and PMID: 28349420). We have added references to those detailed descriptions in the revised manuscript.

7. Would on-target or off-target activity be missed if it overlaps with ssDNA regions associated with endogenous cellular transcription? The strength of CasKAS is that it relies on a novel mechanism of detection, unwound ssDNA, a subset of which in the cell would be associated with CRISPR-Cas activity. A weakness is that it may be challenging to detect true signal if it overlaps with background. For example, it is easy to imagine a more downstream target in the VEGFA gene at the promotor-exon boundary where the on-target peak might be merged with endogenous ssDNA signal. Furthermore, an even-weaker off-target signal might be even more difficult to detect if overlapping.

This is indeed a possible limitation of the *in vivo* CasKAS approach. However, most of the genome is not strongly single-stranded the way active promoters are, and for the purpose of off-target identification (as opposed to tracking their *in vivo* dynamics), the chromatin context is removed, as we carry out the assay on purified DNA.
XXX We have done an experiment with a guide on top of the promoter XXX

8. How does the guanine-frequency in targets and off-target sites affect the ability to detect? Can CasKAS detect targets and off-target sites that contain low numbers of guanine bases equally well as targets that have many guanine bases? The read coverage distribution binned by number of guanine bases in the off-targets would be one way to begin to answer this question.

The requirement that a G is present is a limitation of CasKAS that we ourselves acknowledge. On the other hand, kethoxal labeling is highly efficient, so is the subsequent click chemistry reaction and then the biotin pulldown, this a single G is likely sufficient for capture of a ssDNA bubble.

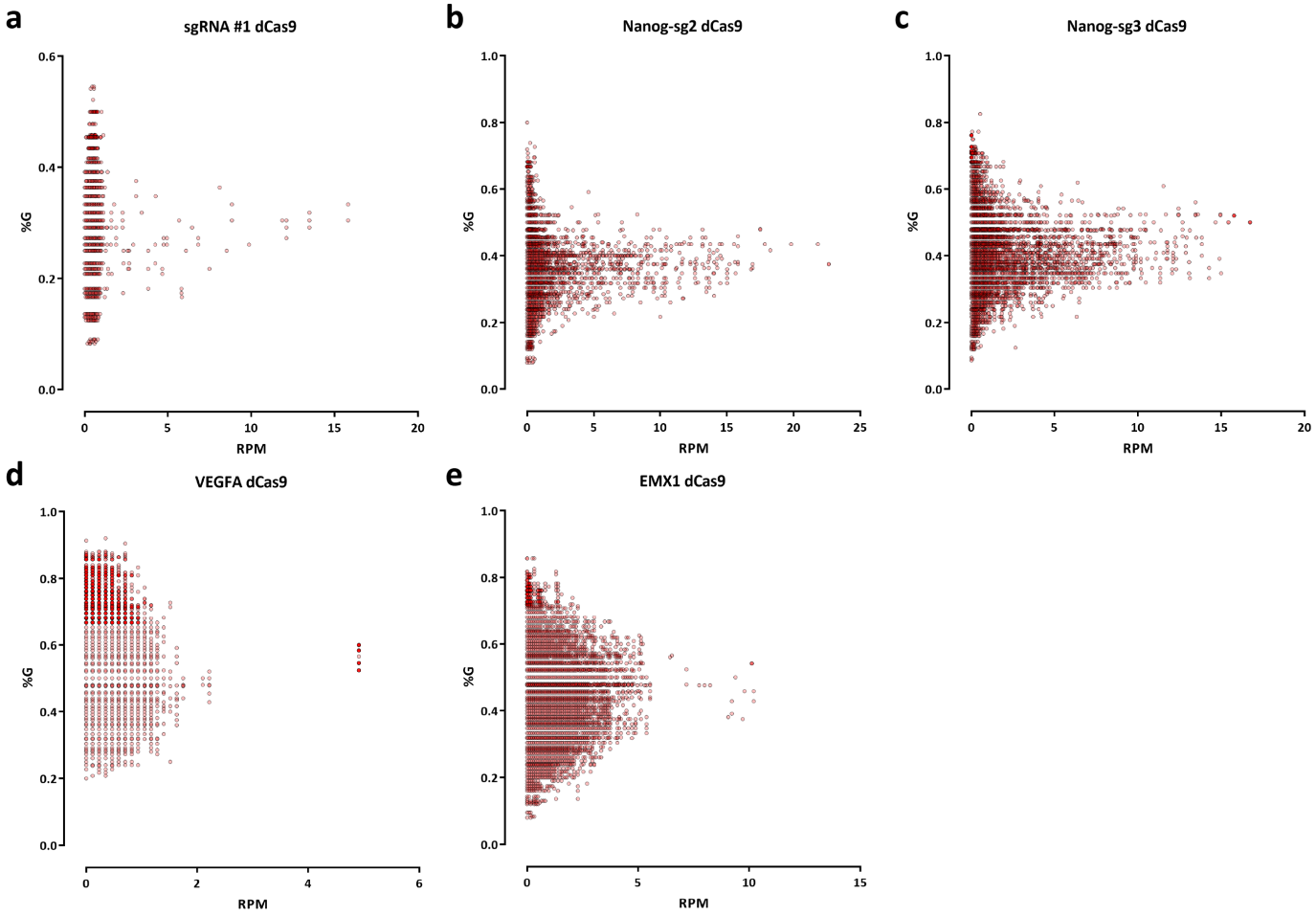
We have compared the G content to the levels of CasKAS and do not find strong correlation between the two. This analysis has been included in the supplement of the revised manuscript and also here as Response To Reviewers Figure 1

9. Why is there an asymmetry in the read coverage surrounding the sgRNA target site? There must be a simple explanation for why the read coverage ends sharply close to the target sequence even for dCas9, but it is not clear from reading the text why that is.

Strand asymmetry is a standard and highly useful feature of *-seq occupancy assays, such as ChIP-seq. It arises because all fragments originate saround the occupancy site, which means that forward-strand reads are located to the left of it and reverse-strand reads are located to the right, as all enriched fragments have to by definition contain the occupancy site. We have expanded that portion of text and provided more references to early ChIP-seq papers to better clarify that point for readers who might not be familiar with the ChIP-seq literature.

Minor 1. A more balanced discussion of both advantages and limitations of earlier methods would be appropriate. For example, discussing the sensitivity strengths of the in vitro methods, the high validation rate of some of the cellular methods, and how some methods capture a broad range of off-target activity that may occur over time and others provide a snapshot of Cas9 activity at a particular moment in time, the scalability of some of the methods, etc.

We thank the reviewer for the suggestion and have expanded the text accordingly. As mentioned elsewhere in the response to reviewers, the strength of CasKAS done *in vitro* is that it can provide a maximally expansive set of off-targets that are in the same time reliable and with minimum off targets.



Response to Reviewers Figure 1: Absence of strong correlation between the number of G nucleotides in a sgRNA off-target site and CasKAS signal. Highly enriched off-target sites do not show a strong preference for containing more G nucleotides than other predicted off-target sites.

Minor 2. Are the other strong peaks shown in Fig. 1e likely real off-target activity or background?

We do not know the answer to this question at the moment. We generally see more "peaks" with active Cas9 than with dCas9, which might be suggestive of lower specificity or occupancy for Cas9 that is being picked up as a ssDNA signal.

Minor 3. For figure panel 1c, it would be helpful to also see an appropriate negative control.

We understand the reviewer's wish, but an example of such a negative control is already provided in the supplement, while main figures need to be kept as concise as possible, and in our opinion the figure is better off as it is.

Minor 4. The scale for Fig. 1c-f and 1i for read coverage is not shown, making it difficult to interpret this figure

We have added the scales.

Minor 5. Fig. 1f is difficult to interpret. Is there only clear activity at the on-target site?

The figure shows the strand-specific cut profiles around the on-target site, in order to illustrate what these profiles look like. It is not the goal of this panel to show global off-target identification.

> Minor 6. It might be helpful to be more specific about the type of Cas9 specificity that is being analyzed in the context of Fig. 2e. Is this Cas9 unwinding specificity (distinct from Cas9 cleavage or Cas9 binding specificity)?

When using dCas9, CasKAS is mapping the formation of structures in which the sgRNA invades dsDNA, leading to the formation of a ssDNA structure.

When using active Cas9, CasKAS is mapping both the formation of ssDNA structures, and cleavage.

We have edited the text to better clarify these points

> Minor 7. Indels in a genome editing context usually refers to mutations associated with the nuclease. In the usage in pg. 4 of the manuscript, the authors use the term to also refer to gaps or bulges in pairing between the sgRNA and the target DNA strand. The wording should be clarified to avoid confusion.

We have edited the text to better clarify that point.

> Minor 8. Similarly, In vivo should be clarified to mean in living cells or in cell culture. Typically, in vivo is used to refer to experiments performed in living animals or plants.

We have edited the text accordingly to make that distinction clear.

Reviewer 2:

> This work describes a genome-wide technique to assess Cas9 specificity (termed CasKAS) by explicitly probing for DNA unwinding events mediated by Cas9. The authors apply published KAS-seq (N3-kethoxal) to covalently label unpaired guanines arising from ssDNA formed during Cas9-mediated binding and DNA duplex destabilization. The modified guanines can undergo a click chemistry reaction, pulled down via biotin, and the DNA fragments sequenced to assess Cas9 unwinding locations. This technique is different from other Cas9 off-target approaches already published in that it interrogates the scope of Cas9-mediated DNA unwinding, as opposed to DNA binding, cleavage, or editing.
> While the idea of the approach is innovative, its utility is not demonstrated. It is unclear what theoretical advantages Cas-induced ssDNA mapping would have over other methods, and the data to support CasKAS is preliminary as compared to other off-target approaches. I dont think the authors have proven that their technique provides an impactful advance in assessing Cas9 specificity.

We thank the reviewer for the comments and suggestions.

> 1. The benefits of CasKAS as presented are mostly theoretical but still not entirely clear to me. Does R-loop detection actually provide a major advance over the (spurious) off-targets identified by dCas-ChIP-seq? The manuscript would benefit from a more thorough comparison to other published off-target detection approaches beyond merely stating the number of identified peaks between each method. A benchmarking between rankings of shared off-targets and the impact thereof could be included. dCas ChIP-seq is dismissed as problematic, but CasKAS also requires large numbers of cells and high sequencing depth. Does CasKAS have any real advantages when compared to other off-target identification methods?

We appreciate the reviewer's concern but are confused about several of the points raised.

First, it is not true that *in vivo* CasKAS requires a large number of cells. The experiments presented in the original version of the manuscript were carried out on 400,000 cells. For a dCas9 ChIP-seq to work well, 20M cells would be needed (this is the typical minimum requirement for a transcription factor ChIP-seq; histone marks can be profile successfully with fewer cells, but the properties of Cas9 occupancy are analogous to those of transcription factors, not of the nucleosomes around which DNA is tightly wrapped).

For comparison, here is a direct quote from the DISCOVER-seq paper:

Cell numbers required

For DISCOVER-seq, we recommend using $\sim 10^7$ edited cells as starting material per chromatin immunoprecipitation. The minimum number of cells to be used as starting material per ChIP is 5×10^6 cells.

And even that statement is in fact not correct. We have more extensive experience than anyone else with QC-ing ChIP-seq data due to our work as part of the ENCODE Project spanning many years, and transcription factor ChIP-seq very rarely works well with less than $\sim 10^7$, the only exceptions being extremely tightly bound TFs such as CTCF and NRSF.

Second, CasKAS in fact very much does not require a high sequencing depth, as shown in Figure 1 in the original manuscript. It only requires 10-20M reads, i.e. what is typical for a ChIP-seq experiment. The off-target detection methods that rely on detecting sequencing edits and those that rely on identifying cleavage sites without enrichment are essentially genome resequencing experiments, requiring one to two orders of magnitude deeper sequencing.

Third, as mentioned in response to Reviewer #1, CasKAS is very quick, easy, and cheap. For a dCas9 ChIP-seq experiment, the reverse crosslinking part alone takes more time than the whole CasKAS workflow.

Strand invasion detection does provide more specificity, as shown in our original analysis.

We have carried out a comparison against additional methods

2. The C-score parameter for cutting efficiency relies on the read counts that were used to call peaks. The authors must test putative off-targets using sequencing.

We are not sure we understand the reviewer's question here. The C-score relates the forward and reverse strand profiles, it is not related to peak calling.

As we explain in our reply to Reviewer #1, edits that can be verified by sequencing arise as a result of cleavage. Thus if one is observing the cleavage itself, that is as good an evidence for productive targeting as one can ask for.

We have nevertheless carried out targeted sequencing as requested.

3. A major limitation of this technique appears to be the signal to noise, especially for in vivo contexts. Read counts (in RPM) for most of the loci are extremely low and barely above background (Figure 1k for instance). Using VEGFA as an example, I question the ability to accurately in vivo call peaks independent of cellular processes that generate ssDNA.

We are afraid that the reviewer is misrepresenting what is showin in Figure 1k. The relevant figure to evaluate enrichment is Figure 1i, where strong and robust enrichment is abundantly clear. What is shown in Figure 1k is the 5' ends around the sgRNA target site. Because only the 5' position receives a score when making tracks that way, they will always look "sparse" as opposed to when coverage over the whole read or fragment is shown (in which case $\geq 2 \times 36$ basepairs will receive a score for each fragment, resulting in a major "pile up" in the coverage track).

The VEGFA examples shows the exact opposite of low enrichment, as the Cas9-induced accessibility peak is half the strength of the strong endogenous KAS-seq peak associated with the active promoter of the gene.

We have carried out additional analysis to further demonstrate these points.

4. Only one time point was tested for addition of kethoxal to the cell lysate, 14 hours post-lipofection of the RNP in vivo. Breaks will be formed in that timespan, and some are even resolved, thus potentially biasing the mapping against loci where Cas9 has already bound, and the DNA edited. A time course could be done to assess the optimal experimental intervals.

XXX We have carried out the requested time course XXX

Minor 1. The methods section should be expanded and more detailed. For example, 1 l of recombinant purified Cas9 under In vitro CasKAS is not informative. Additionally, the sequence analysis section of the methods is not currently detailed enough to thoroughly understand the process. For example, what criteria were used for the manual filtering process?

The phrase "1 μL of recombinant purified dCas9" is immediately followed by "(MCLab dCAS9B-200)", i.e. the catalog number of the dCas9 used. We have added a clarification that this particular MCLab dCas9 is sold at a 20μM concentration, and that we thus used 20 pmol of it.

The manual filtering was explained in the text – "[we] manually curated the resulting peak set, excluding peaks not exhibiting the canonical asymmetric read distribution around a fixed point on the two strands", and this kind of analysis is standard in all high-throughput sequencing-based occupancy assays – but we have now added an explicit reference to the supplementary figure illustrating what strand asymmetry means.

> Minor 2. The presentation of data is not entirely informative in many instances. The heat maps in Figure 2c or Supplementary Figure 8 for example, seem mostly blank. Multiple figures of these, lacking any sequence context, do not convey useful information about the technique.

The goal of the technique is to quantify off-target binding – if there is a lot of it, then these heat maps will be "full", if there is very little of it, then they will be "blank". In the cites cases the heat maps are mostly blank because most predicted off-target sites do not show any binding, i.e. CasKAS did its job of identifying and quantifying off-target occupancy.

> Minor 3. A discussion pertaining to the assessment of in vitro Cas9 unwinding from Ivanov IE, et al 2020, PNAS, could be included to corroborate the off-targets observed with CasKAS.

XXXX

> Minor 4. Figure 1h: Error bars not defined.

We have added a clarification to the figure legend.

> Minor 5. Supplementary Figure 13 is introduced after Supplementary Figure 4 in the text.

We thank the reviewer for noticing that discrepancy. We have rearranged the figures in the correct order of referencing.

> Minor 6. Supplementary Figures 9-12: No description of the black bars at bottom of the graphs.

The black bars represent the extent of sequence conservation for a given position within the multiple sequence alignment. We have clarified that in the revised version.