THE EMBO JOURNAL

Manuscript EMBO-2016-95621

Molecular mechanisms that distinguish TFIID housekeeping from regulatable SAGA promoters

Wim de Jonge, Eoghan O'Duibhir, Philip Lijnzaad, Mr. Dik van Leenen, Marian Groot Koerkamp, Patrick Kemmeren and Frank Holstege

Corresponding author: Frank Holstege, UMC Utrecht

Review timeline:

Submission date:	31 August 2016
Editorial Decision:	29 September 2016
Revision received:	18 October 2016
Accepted:	01 November 2016

Editor: Anne Nielsen

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 29 September 2016

Thank you for submitting your manuscript for consideration by the EMBO Journal. It has now been seen by three referees whose comments are shown below.

As you will see from the reports, our referees all express high interest in the findings reported in your manuscript and support publication in The EMBO Journal, pending satisfactory, minor revision. Furthermore, most of the points raised by referees can be addressed via text revisions and additional data analysis.

Given the referees' positive recommendations, I would like to invite you to submit a revised version of the manuscript, addressing the comments of all three reviewers.

REFEREE REPORTS

Referee #1:

Eukaryotic promoters are frequently distinguished by the presence of TATA box and their dependence on SAGA co-activator, on the one had, or by the presence of a TATA-like element and their dependence on TFIID co-activator, on the other hand. The former class mostly comprises promoters of genes that are highly regulatable, while the second class tend to encompass housekeeping genes. Despite considerable amount of work, the molecular mechanisms underlying the different behaviour of these promoters are far from being completely understood. This situation probably stems from the multiplicity of the mechanisms involved, including the identity of the

transcription activators involved.

In this work, Holstege and colleagues investigate the reason for the high responsiveness of the SAGA-dominated genes and the lower one of the TFIID-dominated genes. To avoid any confounding effect of the identity of the transcription activators, the authors searched for genes that were regulated by a single activator but that were either SAGA- or TFIID-dominated. Hsf1-dependent genes fall into that category with about half of them dominated by TFIID or SAGA.

The authors then investigated how responsive these genes were to the removal of Hsf1 by the anchor away method. The authors also took advantage of the slower removal of the factor when it was induced by lower concentrations of the rapamycin inducer. The experiments indicated that the SAGA-dominated genes were more responsive to Hsf1 removal than the TFIID-dominated genes. The organization of the chromatin environment of the two classes of genes differed, the former class having a less well positioned +1 nucleosome and a less marked nucleosome-free region (NFR). It was found that the negative regulator Mot1 acted preferentially on SAGA-dominated genes, suggesting that it was one of the important factors explaining their responsiveness. Mnase-seq also showed that the +1 nucleosome was more susceptible to Mnase in SAGA-dominated genes suggesting that it is less strongly positioned that that of TFIID-dominated genes.

Altogether the experiments support the conclusion that Mot1 action and nucleosome organization "explain the fundamental differences between housekeeping and regulatable genes". It also nicely shows that a single activator "can elicit different response dependent on core promoter class".

The experiments are well performed and use state of the art methods. The experiments are repeated several times which allowed the authors to perform thorough statistical analyses of their data, fully supporting their conclusion. The manuscript is well written and would be a nice addition to the literature on the mechanisms of gene activation.

My only very minor reservation is that the conclusions drawn here might be relevant to Hsf1 only and not to other transcription activator. The authors might want to add a word of caution in the Discussion section.

Referee #2:

The authors show differences between SAGA-dependent promoters and TFIID-dependent promoters in response to the same activator and propose two molecular mechanisms which contribute to these differences. This is a very interesting manuscript that addresses the important question of whether the core promoter dictates differences in responsiveness to and 'regulatability' by activators, and this is indeed the case. The authors should be applauded for using 4tU-Seq rather than RNA-Seq, as this is more sensitive and directly monitors transcriptional changes. The work is relevant for the large communities working on gene transcription and gene regulation. The text is very well written and the key conclusions are clearly worked out. I strongly recommend publication in the EMBO Journal after the following concerns have been removed:

1. The authors distinguish only between SAGA-dependent/TATA-box promoters and TFIID-dependent/TATA-like promoters, but TFIID does not exclusively function at TATA-like promoters (Basehoar et al, 2004; Huisinga & Pugh, 2004; Rhee & Pugh, 2012), and this should be referred to at least in the discussion.
2. In Fig 2, Fig 4D and 5A-D the authors use 'cor.test' to obtain p-values, while in 4C and Fig 5E the t-test is applied. The reason for this should be given or a consistent test should be applied throughout.
3. It would be nice to known the significance of nucleosomal repositioning on SAGA-dominated promoters (Fig 6). This could be done by calculating the average offset of each measured gene, placing it on the boxplot and making a correlation test.

Referee #3:

This is a very carefully prepared and very clearly written manuscript that describes a series of experiments on a set of 21 heat shock factor 1 (Hsf1) target genes in budding yeast. According to previously published data from the Pugh lab, these genes are nearly evenly divided between those described as SAGA/TBP- and TFIID-controlled, categories previously associated with regulated and constitutive expression, respectively. They consider these genes as a model for understanding how a single transcription factor can interact with two different core promoter elements. This is an important question in the field and has not to my knowledge been addressed previously, at least not in a systematic fashion, since previous studies have not focused on a single activator that is associated with both core promoter types.

The authors first establish that their gene set are indeed Hsf1 targets using an anchor-away approach coupled to qPCR ChIP, following which they measure Hsf1 binding (ChIP-seq read number) as a function of mRNA synthesis change (4-thiouriacil incorporation) 30 min following Hsf1 depletion (under non-inducing conditions). They show that the SAGA dominated genes display a strong correlation between binding and fold-change in transcription, whereas the TFIID set do not. They then use at "slow depletion" protocol (lower rapamycin) and carefully measure fold-change in transcription versus Hsf1 ChIP. This kinetic analysis again supports the idea that the SAGA-dominated genes are more "responsive" (i.e. display a high transcriptional effect for a given difference in Hsp1 binding). They then go on to show that the SAGA-dominated promoters are more sensitive to inhibition by Mot1, a remodeler-like protein that removes TBP from TATA elements. Finally, they show that Hsf1 depletion results in nucleosome shifts at SAGA-dominated promoters, but not at TFIID promoters.

The experiments described here are carefully designed and the data support the authors' conclusions. I have the following specific comments:

1. Much of the argument here is based upon Hsf1 ChIP(-seq) as a measure of binding in vivo. The data in Fig. 2F indicate that there is in general much higher binding at SAGA-dominated promoters, for reasons that are completely unclear, particularly considering that the TFIID promoters seem to have increased nucleosome depletion compared to the SAGA sites. The authors should comment on this at the least, and try to find an explanation for this curious observation.
2. There is an additional assumption that underlies the authors' conclusion that needs to be examined in more detail, and this is related to other possible differences between the two promoter types. For example, it is not clear at all that the authors have examined these promoters for binding of so-called "general regulatory factors" (e.g. Reb1, Abf1, Rap1, etc.), which might be specific to one core promoter type and play a role in the activity of Hsf1. This could be examined by analysis of public databases. In addition, the authors have not discussed the possibility that other sequence motifs (e.g. poly[dA:dT] tracks) might be specific to one type and also play a role in regulation.
3. The authors present very detailed MNase titration data to map nucleosomes at various promoters both in the presence and absence of Hsf1. However, there is little discussion and no explanation for why there should be a very strong Hsf1 effect on nucleosome position at many SAGA genes, but apparently none of the TFIID genes.
4. It is interesting that the authors see two populations of +1 nucleosomes at SSA1 (bottom of p.12). Could this mean that there are 2 cell populations, one with the gene active, the other inactive?
5. I would appear that the TFIID-dominated genes also show increased expression (at least some, and throughout the timecourse) after Mot1 depletion. This is unexpected in light of the results from Zentner where they describe redistribution of TBP from SAGA- to TFIID-dominated genes after Mot1 inactivation. This should be discussed by the authors.

Minor comments:
1. It is unclear why Henikoff et al. (2011) and Kent et al. (2011) are not also referenced with regard MNase-sensitive nucleosomes.

1st Revision - authors' response 18 October 2016

We would like to thank the referees for their time and positive comments. Based on their remarks we have extended the results and discussion. Our responses to the specific points are shown in bold italics below. The changes made to the manuscript are underlined.

Referee #1:
Eukaryotic promoters are frequently distinguished by the presence of TATA box and their dependence on SAGA co-activator, on the one had, or by the presence of a TATA-like element and their dependence on TFIID co-activator, on the other hand. The former class mostly comprises promoters of genes that are highly regulatable, while the second class tend to encompass housekeeping genes. Despite considerable amount of work, the molecular mechanisms underlying the different behaviour of these promoters are far from being completely understood. This situation probably stems from the multiplicity of the mechanisms involved, including the identity of the transcription activators involved.

In this work, Holstege and colleagues investigate the reason for the high responsiveness of the SAGA-dominated genes and the lower one of the TFIID-dominated genes. To avoid any confounding effect of the identity of the transcription activators, the authors searched for genes that were regulated by a single activator but that were either SAGA- or TFIID-dominated. Hsf1-dependent genes fall into that category with about half of them dominated by TFIID or SAGA.

The authors then investigated how responsive these genes were to the removal of Hsf1 by the anchor away method. The authors also took advantage of the slower removal of the factor when it was induced by lower concentrations of the rapamycin inducer. The experiments indicated that the SAGA-dominated genes were more responsive to Hsf1 removal than the TFIID-dominated genes. The organization of the chromatin environment of the two classes of genes differed, the former class having a less well positioned +1 nucleosome and a less marked nucleosome-free region (NFR). It was found that the negative regulator Mot1 acted preferentially on SAGA-dominated genes, suggesting that it was one of the important factors explaining their responsiveness. Mnase-seq also showed that the +1 nucleosome was more susceptible to Mnase in SAGA-dominated genes suggesting that it is less strongly positioned that that of TFIID-dominated genes.

Altogether the experiments support the conclusion that Mot1 action and nucleosome organization "explain the fundamental differences between housekeeping and regulatable genes". It also nicely shows that a single activator "can elicit different response dependent on core promoter class".

The experiments are well performed and use state of the art methods. The experiments are repeated several times which allowed the authors to perform thorough statistical analyses of their data, fully supporting their conclusion. The manuscript is well written and would be a nice addition to the literature on the mechanisms of gene activation.

My only very minor reservation is that the conclusions drawn here might be relevant to Hsf1 only and not to other transcription activator. The authors might want to add a word of caution in the Discussion section.

We agree. A sentence stating this has been added to the discussion.

Referee #2:
The authors show differences between SAGA-dependent promoters and TFIID-dependent promoters in response to the same activator and propose two molecular mechanisms which contribute to these differences. This is a very interesting manuscript that addresses the important question of whether the core promoter dictates differences in responsiveness to and 'regulatability' by activators, and this is indeed the case. The authors should be applauded for using 4tU-Seq rather than RNA-Seq, as this is more sensitive and directly monitors transcriptional changes. The work is relevant for the large communities working on gene transcription and gene regulation. The text is very well written and the key conclusions are clearly worked out. I strongly recommend publication in the EMBO Journal after the following concerns have been removed:

1. The authors distinguish only between SAGA-dependent/TATA-box promoters and TFIID-dependent/TATA-like promoters, but TFIID does not exclusively function at TATA-like promoters (Basehoar et al, 2004; Huisinga & Pugh, 2004; Rhee & Pugh, 2012), and this should be referred to at least in the discussion.

Indeed the difference is not black-and-white. Within the 21 Hsf1 targets there are two SAGA dominated genes that lack a TATA-box, and conversely there are two TFIID dominated genes

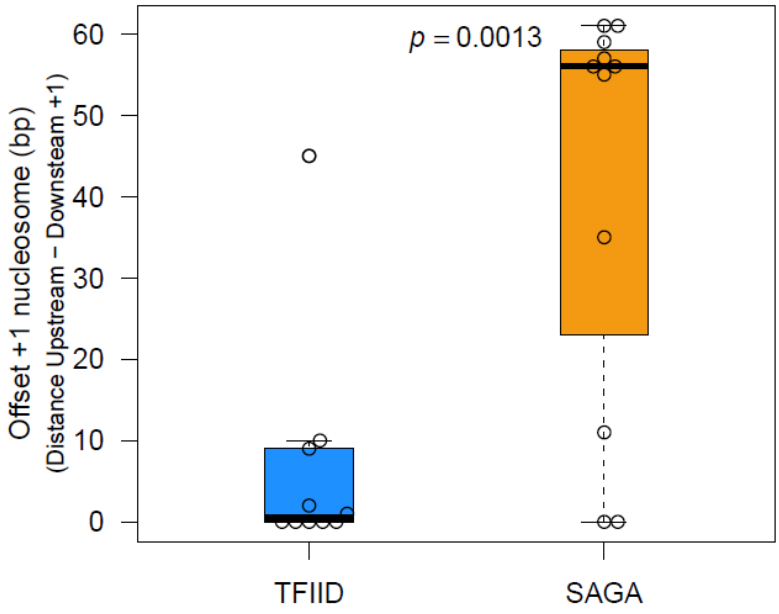
that do have a TATA-box. We now specifically describe this in the discussion.

2. In Fig 2, Fig 4D and 5A-D the authors use 'cor.test' to obtain p-values, while in 4C and Fig 5E the t-test is applied. The reason for this should be given or a consistent test should be applied throughout.

A correlation test is used for assessing correspondence between two different types of data for the same genes (e.g. Hsf1 binding vs expression changes). A t-test is used when evaluating the difference between two groups of genes for the same data (e.g. SAGA genes vs TFIID genes, gene expression response to heatshock). This is now explained in the corresponding figure legends.

3. It would be nice to known the significance of nucleosomal repositioning on SAGA-dominated promoters (Fig 6). This could be done by calculating the average offset of each measured gene, placing it on the boxplot and making a correlation test.

The figure below shows the offset in nucleosome positioning for the two groups of genes as suggested. As expected, the difference is highly significant with a p-value of 0.0013 (t-test).



Referee #3:
This is a very carefully prepared and very clearly written manuscript that describes a series of experiments on a set of 21 heat shock factor 1 (Hsf1) target genes in budding yeast. According to previously published data from the Pugh lab, these genes are nearly evenly divided between those described as SAGA/TBP- and TFIID-controlled, categories previously associated with regulated and constitutive expression, respectively. They consider these genes as a model for understanding how a single transcription factor can interact with two different core promoter elements. This is an important question in the field and has not to my knowledge been addressed previously, at least not in a systematic fashion, since previous studies have not focused on a single activator that is associated with both core promoter types.

The authors first establish that their gene set are indeed Hsf1 targets using an anchor-away approach coupled to qPCR ChIP, following which they measure Hsf1 binding (ChIP-seq read number) as a function of mRNA synthesis change (4-thiouriacil incorporation) 30 min following Hsf1 depletion (under non-inducing conditions). They show that the SAGA dominated genes display a strong correlation between binding and fold-change in transcription, whereas the TFIID set do not. They then use at "slow depletion" protocol (lower rapamycin) and carefully measure fold-change in transcription versus Hsf1 ChIP. This kinetic analysis again supports the idea that the SAGA-dominated genes are more "responsive" (i.e. display a high transcriptional effect for a given difference in Hsp1 binding). They then go on to show that the SAGA-dominated promoters are more

sensitive to inhibition by Mot1, a remodeler-like protein that removes TBP from TATA elements. Finally, they show that Hsf1 depletion results in nucleosome shifts at SAGA-dominated promoters, but not at TFIID promoters.

The experiments described here are carefully designed and the data support the authors' conclusions. I have the following specific comments:
1. Much of the argument here is based upon Hsf1 ChIP(-seq) as a measure of binding in vivo. The data in Fig. 2F indicate that there is in general much higher binding at SAGA-dominated promoters, for reasons that are completely unclear, particularly considering that the TFIID promoters seem to have increased nucleosome depletion compared to the SAGA sites. The authors should comment on this at the least, and try to find an explanation for this curious observation.

The difference is not dramatic. On average there is a two-fold higher degree of Hsf1 binding on SAGA promoters. This is not statistically significant (p=0.07, t-test). We nevertheless used the motif discovery tool MEME to investigate whether differences in the binding motif for Hsf1 between the SAGA and TFIID dominated genes may underlie the two-fold binding difference. The identified motifs were highly similar for both promoter classes. No significant differential flanking motifs were found either. In addition we also looked for motif differences between the strongest Hsf1 binders and the weakest. Again, we found no significant motif difference that could explain a difference in binding. Differences in the number of motifs also don't explain the two-fold difference. One drawback to these analyses is the relatively low number of promoters being analyzed. We now comment on this in the results.

2. There is an additional assumption that underlies the authors' conclusion that needs to be examined in more detail, and this is related to other possible differences between the two promoter types. For example, it is not clear at all that the authors have examined these promoters for binding of so-called "general regulatory factors" (e.g. Reb1, Abf1, Rap1, etc.), which might be specific to one core promoter type and play a role in the activity of Hsf1. This could be examined by analysis of public databases. In addition, the authors have not discussed the possibility that other sequence motifs (e.g. poly[dA:dT] tracks) might be specific to one type and also play a role in regulation.

We had looked at binding of these "general regulatory factors" (GRFs) using several publicly available datasets (Reb1 and Abf1, Kasinathan et al. Nature Methods 2014; Reb1 and Rap1, Rhee and Pugh Cell 2011). We found that within the Hsf1 targets the ratio of SAGA to TFIID dominated promoters bound by any GRF is almost 50/50: Abf1 (2:2), Rap1 (2:2), Reb1 (4:3) (or (5:4) depending on the dataset). In total 9 out of 11 SAGA dominated genes and 6 out of 10 TFIID dominated genes had binding of any GRF. It is therefore unlikely that the GRFs play a role in the difference between the Hsf1 TFIID and SAGA genes.
We now describe this in the results.

Similarly, we have investigated a differential presence of various length poly[dA:dT] stretches but could find no clear difference.
We now describe this in the results.

3. The authors present very detailed MNase titration data to map nucleosomes at various promoters both in the presence and absence of Hsf1. However, there is little discussion and no explanation for why there should be a very strong Hsf1 effect on nucleosome position at many SAGA genes, but apparently none of the TFIID genes.

One possibility is that TFIID promoters have an intrinsic "ability" to form nucleosome depleted regions (NDRs), while for SAGA dominated genes, remodelers are continuously needed. This fits with the larger NDRs found on TFIID promoters.
We now include this in the discussion.

4. It is interesting that the authors see two populations of +1 nucleosomes at SSA1 (bottom of p.12). Could this mean that there are 2 cell populations, one with the gene active, the other inactive?

Yes, this is exactly what we think is going on and we now specifically state this in the results.

5. I would appear that the TFIID-dominated genes also show increased expression (at least some,

and throughout the timecourse) after Mot1 depletion. This is unexpected in light of the results from Zentner where they describe redistribution of TBP from SAGA- to TFIID-dominated genes after Mot1 inactivation. This should be discussed by the authors.

It is is not completely unexpected to observe some TFIID genes with increased expression upon Mot1-depletion (Figure 5E). In Zentner et al. TBP redistribution was monitored upon temperature-sensitive Mot1 inactivation. Zentner et al. observed that genes with increased TBP binding upon Mot1 inactivation, tend to have a TATA box. This does not mean that every single gene that has increased TBP binding is a TATA-box containing / SAGA dominated gene. If we look at the same top 500 sites of TBP increase as Zentner did, approximately 50% have a TATA-box (based on Rhee and Pugh, 2012). This is an enrichment compared to the ~20% genome-wide average, but also still means that half of the genes with increased TBP upon Mot1 inactivation do not have a TATA-box. It is therefore not unexpected to observe a slight increase in expression of a few TFIID genes in our Mot1 depletion experiment (Figure 5E). This relates very much to the first comment of referee#2 and we now specifically describe this in the discussion.

Minor comments:
1. It is unclear why Henikoff et al. (2011) and Kent et al. (2011) are not also referenced with regard MNase-sensitive nucleosomes.
These references have now been added.

Additional changes:
We have split figure 6 into figure 6 and figure 7 because the original figure 6 did not fit the format for a single page.

2nd Editorial Decision 01 November 2016

Thank you for submitting a revised version of your manuscript. It has now been seen by one of the original referees and this person's comments are shown below. As you will see the referee finds that all criticisms have been sufficiently addressed and recommends the manuscript for publication. I am therefore happy to inform you that your study has been accepted in The EMBO Journal.

REFEREE REPORT

Referee #3:

The authors have satisfactorily replied to all of our comments and the manuscript is now suitable for publication in The EMBO Journal

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: F.C.P. Holstege
Journal Submitted to: EMBO Journal
Manuscript Number: 95621

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	For the microarrays, samples were measured as biological replicates in dye-swap. The ChIP-seq material was created from several independent IPs.The ChIP-qPCR, ChIP-chip and MNase-seq were done with at least 3 independent biological replicates.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Only one gene was included from the analysis, because of high cross-hybridization on our microarrays.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes, see figure legends and materials and methods.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes, everywhere there is a significant difference in using a t-test, both groups are normally distributed (inspected using a shapiro-test).
Is there an estimate of variation within each group of data?	For each comparison all indivial datapoints are shown.
Is the variance similar between the groups that are being statistically compared?	See above

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Yes, see figure 1G-H no tag control.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	All sequencing and microarray data has been deposited to GEO: GSE81481.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	See above
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	See above
22. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

23. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA