THE EMBO JOURNAL

Integrated requirement of non-specific and sequence-specific DNA binding in MYC-driven transcription

Paola Pellanda, Mattia Dalsass, Marco Filipuzzi, Alessia Loffreda, Alessandro Verrecchia, Virginia Castillo Cano, Hugo Thabussot, Mirko Doni, Marco Morelli, Laura Soucek, Theresia Kress, Davide Mazza, Marina Mapelli, Marie-Eve Beaulieu, Bruno Amati, and Arianna Sabò
DOI: 10.15252/embj.2020105464

Corresponding author(s): Bruno Amati (bruno.amati@ieo.it) , Arianna Sabò (arianna.sabo@ieo.it)

Review Timeline:

Submission Date:	1st May 20
Editorial Decision:	5th Jun 20
Revision Received:	7th Jan 21
Editorial Decision:	4th Feb 21
Revision Received:	15th Feb 21
Accepted:	24th Feb 21

Editor: Stefanie Boehm

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. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

5th Jun 2020

Re: EMBOJ-2020-105464
Integrated requirement of non-specific and sequence-specific DNA binding in MYC-driven transcription

Dear Dr. Amati,

Thank you for submitting your manuscript for consideration by The EMBO Journal. Please also excuse the delay in communicating this decision to you, which was due to delayed referee reports on account of the current pandemic, as well as the high number of new submission we are currently receiving. We have now however received three referee reports on your study, which are included below for your information.

As you will see, the reviewers are overall positive and acknowledge the interest of the field in the topic and the study. Nonetheless they also raise some concerns that would need to be addressed in a revised manuscript. In particular, both referee #1 and #3 find that further characterization of the Myc mutants used, would be needed, including determining the protein half-life and alternative assessment of binding affinity (ref#1 point 1; ref#3 points 1, 4, 5, 6). These issues should be experimentally addressed. In addition, referee #2 points out that the conclusion that unspecific binding is a prerequisite for specific binding is not formally shown (point 1, 2). Moreover, s/he raises questions regarding the residual binding of MycHEA and its dominant negative effect on wild-type Myc (point 4), which are similar to those of referee #1 (points 2, 7). These issues should also be resolved in the revised manuscript by additional experimental data or, when applicable, discussion and textual revision. The remaining points of the referees can largely be addressed by additional analyses of available data or further discussion and clarification, and should be carefully responded to. In particular, further discussing the results of the study in context of the models for Myc transcriptional activation will be important (i.e. ref#1 point 8).

Please note that it is our policy to allow only a single round of major revision. We realize that lab work worldwide is currently affected by the COVID-19/SARS-CoV-2 pandemic and that an experimental revision may be delayed. We can extend the revision time when needed, and we have extended our 'scooping protection policy' to cover the period required for a full revision. However, it is nonetheless important to clarify any questions and concerns at this stage and we encourage you to discuss a revision plan and any potential issues you may foresee as soon as possible.

Please also feel free to contact me should you have any other further questions. Thank you for the opportunity to consider your work for publication, I look forward to receiving your revised manuscript.

Kind regards,

Stefanie Boehm
Stefanie Boehm
Editor
The EMBO Journal

Referee #1:

Using several solved structure of MYC, Pellanda et al generate two mutations in the basic region of murine Myc that are predicted to eliminate binding to DNA at either the phosphodiester backbone (MycRA) or the E box motif (MycHEA). The authors then asses the effects of these mutations by expressing the Myc mutants in different cell systems systems. The RA mutation impairs Myc's ability to bind chromatin, promote transcription, and drive cellular processes associated with Myc indicating that non-specific interaction with DNA is essential for Myc activity. The HEA mutation is more nuanced. When expressed at high levels it binds more genomic sites, increases transcription, and has a dominant-negative effect on WT Myc in cells. At endogenous levels MycHEA engages fewer genomic sites (especially with canonical E boxes), is crippled for cell proliferation, and prefers a novel E box derivative (CACGTC).

Based on their observations, the authors propose a model in which Myc binds to open chromatin in a sequence independent fashion followed by localization to E box elements that stabilize Myc and drive transcription. The authors' data sides with a model in which Myc activates specific genes rather than acting as a general transcriptional amplifier. The manuscript is interesting and the data would be of significant interest to the Myc community and likely to transcription regulation in general, but some issues need to be addressed.

Since the cellular levels of Myc correlate with binding to target genes, the manuscript could benefit from a more thorough characterization of the Myc mutations to insure that the only difference between the three versions of Myc is their affinity for E box motifs.
Do these mutations alter the stability of Myc?
Can the authors perform an assay that is more sensitive and quantitative than gel shift to determine the binding affinities to a canonical E box?
The authors show that MycRA and MycHEA engage Max at levels similar to the wild type protein (Fig EV1-C), but this is in an over-expression system. The authors should perform co-precipitation experiments for binding to Max ideally in cb9 cells, but at a minimum in the 3T9 system.

Can the authors provide an explanation as to why MycHEA recognizes partial E-box motifs by ChIP (Fig 1J) but is unable to bind an Ebox by gel shift (EV1-D)? If the author's suggestion that Max is responsible for recognition of an E-box half-site is correct, then why doesn't MycRA also recognize partial E box elements as both mutants still retain binding to Max?

In Fig 2 changes in transcription are consistent with the presence of a canonical Ebox for MycWT or CACGTC for MycHEA in 3T9 cells. But do these transcriptional changes correlate with the distribution of MycHEA on chromatin from the 3T9 ChIPseq? Additionally, the authors should perform qPCR on a selection of genes in cb9 cells and relate this to the cb9 ChIPseq results to determine whether the changes in transcription persist when MycHEA and MycRA are expressed at endogenous levels.

The immunoblot for Myc expression in cb9 cells without Dox (Fig EV3-B) is not convincing and should be performed again to show levels of MycWT.

How do the CACGTC sites bound by MycHEA in cb9 cells correlate with chromatin accessibility? Can the authors compare their cb9 ChIPseq to the published H3K4 and H3K27 data sets shown in Fig EV1-A? Is there anything else special about these sites?

It is interesting that the ability of Myc to bind chromatin is lost with the RA mutation even though MycRA should retain affinity for E boxes. Is it possible that non-specific contacts to DNA are required for Myc to stably bind DNA regardless of local scanning? In other words, minimal Myc binding at an E box requires the sum of non-specific and specific interactions with DNA.

Many transcription factors other than MYC recognize E box elements, yet MycHEA behaves as a dominant negative only in the presence of wild type Myc. Can the authors propose a mechanism for this level of specificity?

Can the authors further elaborate on how their data supports a model of specific gene activation by Myc vs the general transcriptional amplifier model? This is an important conclusion given the controversy between these two models in the literature.

Referee #2:

The authors Pellanda & Dalsass et al. describe in their manuscript entitled "Integrated requirement of non-specific and sequence specific DNA binding in Myc-driven transcription" that non-specific binding of Myc to chromatin is a pre-requisite for subsequent sequence-specific E-box binding, and that sequence-specific binding is required for gene activation of specific transcriptional programs. The authors demonstrate that despite the ability of Myc mutants lacking E-box recognition to exhibit non-specific interactions with chromatin, that these interactions lack specificity and the ability to contribute to E-box associated gene activation and normal proliferative effects observed with wild-type Myc. Overall, the study is very interesting and sheds light on the clouded field of Myc DNA binding and transcriptional regulation.

The authors utilize Myc mutants unable to bind to DNA (MYCRA), and a MYCHEA mutant that lacks affinity for the canonical Myc E-box sites. These tools are powerful and informative. They also take advantage of cells engineered to delete endogenous Myc to allow for lower levels of expression of the mutants versus WT. At lower levels, the MYCHEA mutant has reduced affinity for DNA in general and some selection for a non-canonical E-box site. This results in a change in gene activation relative to WT Myc and the inability to support proliferation. These results support a role for sequence-specific binding in Myc transactivation activity.

Major concerns:

1) One of the authors' major claims, repeatedly stated throughout the manuscript, is that non-specific DNA binding is required for the subsequent sequence-specific binding to E-box sites. Even the title calls this an "integrated requirement". However, based on the two mutants used, it is not clear how this claim can be made since the mutants either lack DNA binding, or lack E-box specific binding. There is no mutant that can bind to E-box sites but not non-specifically to DNA (which would likely be difficult to create). This claim needs clarification to either keep the claim or to revise the interpretation.

2) If the claim is kept, mechanistic insights for how non-specific binding enhances the specific binding would be important.
3) In figure 2B, the authors show 4 and 8 hr time points and differences in genes regulated by WT vs the HEA and RA Myc mutants. Can the authors identify a pattern of increasing gene expression over time with WT (or some temporal pattern) that is not apparent in the HEA mutant? This might shed light on transcriptional regulation mechanisms and gene programs lost with non-specific binding.
4) There does appear to be some residual binding of MycHEA to E-box sites that the authors attribute to half sequence recognition via the heterodimer MAX. What about tetramers with endogenous MycWT? Is this still observed with the deletion of endogenous Myc? Likewise, the dominant negative activity of MycHEA with endogenous Myc present could reflect the HEA mutant reducing affinity of endogenous Myc in a tetramer with it from canonical sites needed for gene activation and proliferation. Further discussion would be useful.
5) The data in EV5C showing that the MycHEA mutant can repress Myc regulated genes is striking and suggests that the MycHEA mutant antagonistically occupies these sites without transcriptional activation for some reason. But, in lines with the above comment, if this mutant does not specifically bind to the Ebox sites, how is it doing this? It would be interesting to know from the ChIP-seq data if the MycHEA mutant truly occupies these sites, which could help with mechanism and would support that Myc binding to specific DNA motifs is permissive to transcriptional activation, as opposed to general occupancy by Myc at the enhancer/promoter regions.
6) In Figure EV2B the authors show greater invasion of chromatin with the MycERHEA mutant, and wish to determine whether this is due to the aberrant overexpression of the Myc mutants. In figure EV3 the authors utilize alteration of the endogenous alleles of cb9 cells with the HEA modification, and compare this to the enforced retroviral expression of the mutants with silenced endogenous expression of Myc. There is significantly less binding with the HEA mutant without endogenous Myc and even less with the HEA knock-in. Further discussion of this relative to what is shown in Fig 1 with greater general binding with the HEA mutant would be helpful as the prominence of Fig 1 leaves the reader expecting that the HEA mutant has more widespread binding.
7) The authors claim that sequence-specific DNA binding activates gene expression at three distinct levels: stabilization, positioning, and promotion of its transcriptional activity. It is not clear that these are distinct mechanisms, please clarify.
8) If sequence-specific DNA binding stabilizes Myc to DNA then why does the WT and HEA both have equivalent decrease in mobility on DNA? Please discuss this further.
9) The HEA retroviral expression and HEA knock-in without endogenous Myc show substantially reduced general DNA binding activity (EV3C). Especially compared to the WT knock-in, which is expressed even lower and looks to be barely detectable in EV3B. Further discussion on this and the role of endogenous Myc in the HEA mutant's DNA binding in ChIPseq experiments with endogenous Myc and in the interpretation of the HEA data showing loss of Myc gene programs and proliferation is important.
10) Further, comparing proliferation between the MycWT and HEA with loss of endogenous Myc, which shows HEA reduced proliferation activity, please discuss whether the overall reduction in its DNA binding shown in EV3C could be an explanation.
11) Page 8 line 10, down regulated genes by WT or HEA lacked E-box sites, but the next sentence says the HEA mutant represses Myc targets that have E-box sites. Please clarify.

Minor points:
1. It would be interesting to speculate as to why the HEA mutant is expressed at higher levels compared to WT in the cb9 knock-ins.
2. On page 5, line 3, it is not clear what "binding hierarchy" is referring to, please clarify.

3. On page 8 lines 3 and 5 reference figures S6A and S6B, although there are no such figures.

Referee #3:

This is a fascinating and important paper that addresses previously unexplored yet fundamental activities of the Myc protein. The authors set out to address the question of how non-specific and specific DNA-binding activities of Myc affect its genomic binding and transcriptional regulation. By mutating residues that contact either the DNA backbone (MYC RA) or the E-box consensus motif (MYC HEA), the authors show that non-specific DNA binding is required for Myc to bind to the genome, while specific sequence recognition is essential for maintaining Myc's normal transcriptional profiles and biological activity. These findings have relevance for interpreting ChIP_Seq data for MYC as well as for understanding how Myc influences expression of its target genes. The data presented are in general thorough and compelling although several issues should be addressed.

1. In Fig. EV1D, it would be important to show whether the Myc mutants bind probes non-canonical E-box probes in vitro? Why does MycHEA display some binding to E-box DNA in vivo but not in vitro?
2. In Fig. 1A and EV2B, it appears that MycERHEA binds genomic DNA more strongly and widely in vivo than MycERWT? What is the explanation for this effect?
3. In Fig. 1C, the differences in peak intensities among canonical E-box, variants and motif-free regions for MycHEA appear similar to MycWT. Does this weaken the conclusion that preferential binding of E-boxes is lost in MycHEA expressing cells.
4. Fig. EV3B and 3C are a somewhat confusing. Is retroviral expression of Myc variants considered as overexpression like the MycER system? If so, why is there less binding with MycHEA expression compared to MycWT (even though MycHEA expression is higher), which is opposite from the MycER results. Interestingly, cb9-mycHEA is more highly expressed than cb9-mycWT, but still binds less. Is it because MycHEA requires WT Myc to be able to better bind to DNA, or it needs to reach a certain threshold of expression to be able to bind DNA better than MycWT? It would be helpful to have Western blots comparing expression levels of Myc variants in all three systems (MycER, retroviral expression, and cb9-myc).
5. In Fig. 1F and EV4D, is the increase in residence time with HEA significant compared to WT? If so, why?
6. Related to the above points - although in several immunoblots the levels of the WT and mutant Myc are similar it would be reasonable for the authors to actually determine the half-lives of the WT and mutant proteins in the cells used here. Myc proteolysis has been suggested to be linked to its transcriptional activity. If the mutant proteins were stabilized then it could explain changes in protein mobility and possibly other aspects of the behavior of these proteins.
7. Is CACGTC bound motif more enriched in cb9-mycHEA cells than in MycERHEA cells? Is the enrichment for degenerate AC-rich motifs lost in the cb9-mycHEA cells?
8. In Fig. 2F and EV5B, is CACGTG motif significantly enriched in down-regulated genes in MycERHEA cells? If so it would be consistent with the dominant negative effect of MycHEA.
9. What specific growth related and metabolic pathways are positively enriched in MycHEA cells? Does MycHEA inhibit cell proliferation mainly by failure to activate Myc target genes or by activation of a different subset of growth-suppressive genes?

Minor:
1. In Fig. EV1C, is the shift of MycRA significant compared to MycWT?
2. On page 8, line 5-8, Fig. S6A or Fig. S6B are not included in the manuscript.
3. What's the point of showing the "spike in" in EV1?

EMBOJ-2020-105464R

Authors' Rebuttal

Authors' general points for all Referees:

We thank the Referees for their thorough assessment and constructive remarks. We believe that these have contributed to significantly strengthen and clarify the concepts made in our work.

Prior to a point-by-point rebuttal to the comments of the Referees, we list here a series of points (Authors' points 1-4) that we deem of general relevance.

- ***Authors' point 1: biochemical characterization of the mutants.***

A common remark by the referees regarded the need to provide a fuller biochemical characterization of the dimerization and DNA-binding properties of our Myc mutant proteins. We fully agree with this critique, and in particular on the importance of making sure that the mutations introduced in the basic region of Myc impair the DNA binding, but not dimerization activities of the protein (as already suggested by our original co-IP data). We have now included several new experiments, which we believe thoroughly confirm this point.

(i.) Dimerization:

*We have now used Circular dichroism (CD) spectroscopy to address the dimerization activities of recombinant Myc and Max bHLH-LZ polypeptides. Quoting from our text (p.4), the data " revealed similar helicoidal content and thermal denaturation profiles for all the heterodimeric Myc/Max complexes (**Fig. 1C**) indicating equivalent dimerization properties of Myc^{WT}, Myc^{RA} and Myc^{HEA} with Max."*

As a complement, we have also extended our co-IP analyses. Besides the original co-IP data in transiently transfected 293T cells (now in Fig. EV1F), we now produce similar data in $3T9^{MycER}$ and $cb9^{\Delta myc}$ cells expressing the various mutant forms (Fig. EV1D, E).

Altogether, our data clearly confirm that, as expected based on the structure (Fig. 1A) as well as on previous functional data (in particular for Myc^{RA}; Amati et al. 1992), the RA and HEA mutations do not significantly affect dimerization with Max.

(ii.) DNA binding:

The reviewers correctly remarked that, while binding DNA in vivo, as judged by ChIP-seq, the Myc^{HEA} mutant did not show DNA-binding activity in the gel shift (EMSA) assay. We have repeated the EMSA experiments with new proteins preparations: unlike in our original experiments, we have now detected binding of Myc^{HEA}/Max dimers to an E-box (CACGTG) oligonucleotide (Fig. 1D) and shown that this occurs with reduced affinity relative to binding Myc^{WT}/Max (Fig. 1E-F). On the other hand, the new data confirmed that Myc^{RA}/Max dimers show full loss of DNA binding activity. We kindly point the referees to our revised manuscript for a full interpretation of the data (p. 4, paragraph starting with "In order to characterize the dimerization and DNA-binding activities…").

In the new EMSA experiments (Fig. 1D-H), we used the same His6-tagged proteins as for the CD analyses of Fig. 1C. Most importantly, a new preparation of non-tagged proteins (analogous to those used in our original experiments) confirmed the same results, as shown here for the referees (Fig. R1). Hence, our new results are not a mere consequence of the addition of the His6 tag.

We note here that the failure to detect the DNA-binding activity of Myc^{HEA}/Max in our original EMSA experiments, while detectable in the new ones, is most likely attributable to differences in the quality and/or purity of our recombinant protein preparations. In particular, although an initial assessment by MALDI and SDS-PAGE showed acceptable protein quality for EMSAs, the same control performed a posteriori revealed that our original batch of the Myc^{HEA} protein had suffered from partial degradation at some point during the analysis, possibly due to sample contamination. With the new batches of either non-tagged or His6-tagged constructs, we verified that proteins integrity was maintained at onset, during and after the experimental analyses. That both the non-tagged and tagged versions now provide the same results (Fig. R1), further strengthens the validity of the new data.

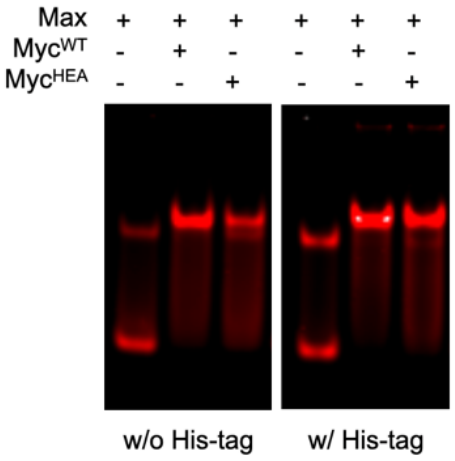

Figure R1: *Independent preparations of non-tagged (left) and His6-tagged (right) Myc and Max peptides provide consistent results in EMSA with a labeled CACGTG probe (for details, see Fig. 1D in the manuscript).*

It is also noteworthy that detecting DNA-binding activity in the gel shift assay requires formation of stable protein-DNA complexes that must resist separation during sample processing and gel electrophoresis: thus, unstable complexes with high off-rates would dissociate and be lost in these experiments. In line with these considerations, early gel shift experiments with mutant probes suggested that the binding of wild-type Myc/Max to E-box oligonucleotides was affected by mutations that alter half of the E-box motif (e. g. Littlewood et al. 1992 Oncogene 7, 1783): for this reason, we were not surprised by the apparent failure of Myc^{HEA}/Max to retard the CACGTG probe in our original experiments – a premature conclusion in retrospect, for which we want to acknowledge the appropriate critique from the referees.

As presented in our first submission, careful analysis of our ChIP-seq data revealed that Myc^{HEA} not only lost affinity for the canonical CACGTG motif, but aberrantly recognized the variant motif CACGTC. We now present an EMSA experiments that include titration of Myc protein concentrations (Fig. 1E), as well as competition with various unlabeled oligonucleotides Fig. 1F-H). Altogether, the new data confirm the altered specificity of Myc^{HEA}/Max relative to Myc^{WT}/Max dimers. Please refer to the relevant parts of the text for a full description (p.4 from "Two experiments were performed to assess the relative DNA-binding efficiencies of Myc^{WT}/Max and Myc^{HEA}/Max dimers…" and p. 9, from "This change in binding specificity was confirmed in a competitive EMSA assay…").

In a general manner, EMSA or other in vitro assays based on the binding of recombinant proteins to short oligonucleotides should be interpreted with caution when assessing weak residual interactions of transcription factors with half- or degenerate sites. As exemplified in our work, such interactions may instead be recognizable by computational analysis of ChIP-seq profiles: indeed, the propensity of Myc^{HEA}/Max dimers to associate with the variant site CACGTC was unraveled by the ChIP-seq analysis (Fig. 3E-H, Fig. EV2C-E), and subsequently confirmed by our competitive EMSA experiments (Fig. 1F-H). Most importantly however, and as explained in our manuscript (p. 9), this "did not amount to a full subversion of DNA-binding specificity": please refer to the text for further detail.

- ***Authors' point 2. General evidence for the integrated and sequential requirement of non-specific DNA binding and sequence recognition in Myc activity.***

Altogether, the ChIP-seq data on our mutant Myc proteins (both as MycER and full-length Myc) provided two essential observations: first, the Myc^{RA} mutant, lacking DNA-backbone interactions, shows an overall loss of DNA binding (Fig. 3A; Fig. EV3A-C); second, while unable to recognize the E-box, the Myc^{HEA} mutant still binds DNA non-specifically, and shows largely unchanged genomic distributions along active chromatin in comparison with Myc^{WT}.

The above data are complemented by two distinct observations, now grouped into a new section subtitled "Non-specific DNA binding restrains free diffusion of MYC in the nucleoplasm". First, as already reported in our original submission, single-molecule tracking experiments (now in Fig. 4B, C and Fig. EV4) showed that Myc^{RA}, unlike Myc^{HEA}, had increased intra-nuclear mobility relative to Myc^{WT}. These results are now complememented by an independent approach, based on subcellular fractionation: quoting from our text (p. 8), '"3T9-MycER cells (grown with OHT) were used to prepare three subcellular fractions (cytoplasm, nucleoplasm and chromatin) and protein distribution analyzed by immunoblotting: MycER^{WT} and

*MycER^{HEA} showed roughly equal proportions of the protein in the three fractions, while MycER^{RA} was essentially lost from chromatin (**Fig. 4A; Fig. EV4A**).”*

Hence, added our ChIP-seq analyses, the mobility and fractionation data lend strong support to the sequential binding model put forward in our work: Myc^{RA} is unable to establish the first contact with the DNA backbone, which obviously would also prevent sequence recognition. As concluded in our text (p.8), “non-specific DNA binding is required for the initial engagement of Myc onto accessible genomic regions, restricting its free diffusion in the nucleoplasm and potentially allowing localized, linear scanning of the DNA sequence.”

Finally, our detailed comparison of DNA recognition and transcriptional regulation by MycER^{WT} and MycER^{HEA} support our conclusion (p.10) “that sequence recognition is essential to establish adequate Myc-activated programs. Most importantly, this step is subsequent to engagement of the factor on active chromatin, mediated by non-specific DNA binding.” These aspects are now explicitly developed in our Results and Discussion sections.

Altogether, the proposed model constitutes the most conservative explanation to unify all of the available data in the field. In this regard, as prompted by the Referees, we have critically revisited alternative interpretations in our Discussion, including the concept of chromatin “invasion” and the “transcriptional amplifier” model (p.12-13).

- ***Authors’ point 3: elimination of the Myc^{HEA} knock-in clones***

*In our original manuscript, we had engineered transgenic tet-Myc fibroblasts in two ways. On one hand, as described in our current text (p.5), we had “inactivated the endogenous c-myc locus by genome editing (cb9$^{\Delta myc}$; **Fig. EV2A**) and transduced the cells with cDNAs encoding Myc^{WT}, Myc^{HEA} or Myc^{RA}, thus leaving only the latter proteins upon tet-Myc shutdown (**Fig. EV1B**)”: the data in this model were confirmed and extended, as will be described further below.*

On the other hand, we had also introduced the HEA mutation into both of the endogenous c-myc alleles (cb9-myc^{HEA}). As correctly pointed out by the referees, the immunoblot analyses presented in our original submission were preliminary, and were thus repeated during revision. In Figure R2A, included here for the referees, we show a unique blot displaying the levels of the retrovirally expressed MycER and full-length Myc proteins (this part is the same as in Fig. EV1C in our revised manuscript) followed by two control cb9-myc^{WT} clones and two knock-in cb9-myc^{HEA} knock-in clones, revealing lower Myc protein accumulation in the latter.

The same result in the four clones is reproduced in Fig. R2B (lanes 1-4), compared with the same cells treated with the proteasome inhibitor MG132 (lanes 5-8): while MG132 stabilized the protein in all of the cultures, as expected, the cb9-myc^{HEA} knock-in cells still showed lower levels than the cb9-myc^{WT} controls. Hence, defective protein accumulation in cb9-myc^{HEA} cells cannot merely be explained by preferential degradation of the Myc^{HEA} protein.

Altogether, while the causes of the defective accumulation of Myc^{HEA} in the knock-in cells remain to be fully clarified, this cannot be simply explained by a lower stability of the mutant protein, in line with the equivalent half-lives of the wild-type and mutant proteins when ectopically expressed, either as full-length Myc or as MycER proteins (Fig. EV2B). We surmise that the introduction of the mutation in cb9-myc^{HEA} cells must be impacting on the synthesis of the protein at a different level (e.g. codon usage, mRNA folding, etc…).

As a corollary, the altered DNA-binding profiles that we had observed by ChIP-seq in cb9-myc^{HEA} cells are no longer interpretable as a primary effect of the HEA mutation, but may merely be due to the lower levels of the Myc^{HEA} relative to Myc^{WT}, when expressed form the endogenous c-myc locus. For this reason, we have decided to remove these data from the paper.

Most importantly, this bears no impact on any of the other experiments presented, with either the Myc^{HEA} of MycER^{HEA} forms, for which all formal controls are provided and the data fully conclusive. Altogether, the removal of the cb9-myc^{HEA} knock-in clones does not detract on any of the conclusions presented in our manuscript.

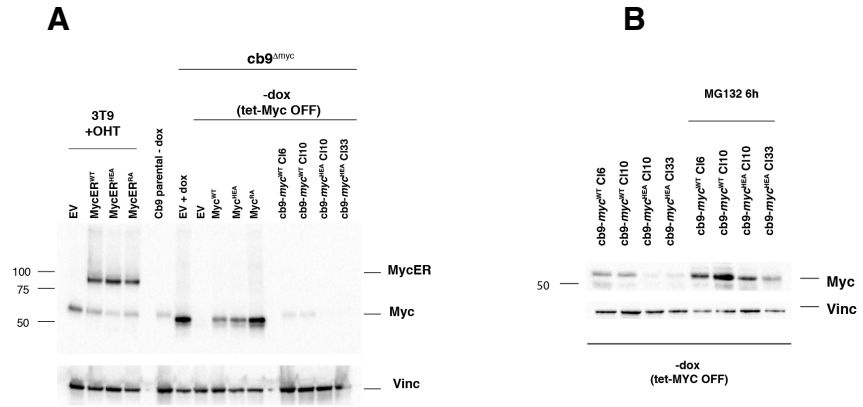

Figure R2: ***A.*** *Immunoblot analysis of 3T9 and cb9$^{\Delta myc}$ cells infected with retroviral vectors expressing the indicated MycER and Myc proteins (the same as in Fig. EV1C), followed by the indicated cb9-myc^{WT} and cb9-myc^{HEA} clones.* ***B.*** *The same clones were treated with the proteasome inhibitor MG132 (10 μM, 6h) prior to immunoblotting. Note that the protein in the cb9-myc^{HEA} clones is stabilized, but not to the levels reached in the cb9-myc^{WT} clones, suggesting that defective accumulation of the Myc^{HEA} protein in those clones is not merely to more effective degradation.*

- ***Authors' point 4: Myc^{HEA} is a dominant-negative mutant***

Another common remark by the Referees regarded the nature and mechanistic basis of the dominant-negative (DN) activity of the Myc^{HEA} mutant. This is indeed an important aspect, which we now describe and interpret with due attention in the manuscript.

The data addressing the biological activities of our Myc mutants in cells have been moved to the beginning of our Results, in the section entitled "Myc^{RA} and Myc^{HEA} are unable to sustain cell proliferation": the DN activities of the HEA mutant (either as MycER or full-length Myc) and its dependency upon endogenous Myc are presented in this context, allowing a clearer description of the main observations. Subsequently, our ChIP-seq and RNA-seq profiles provide a coherent framework, in which to integrate all of the data and interpret the DN activity of Myc^{HEA}, which we do in our Discussion.

Synthetically here, the data show that Myc^{HEA} does not recognize E-box elements, implying that its dominant-negative activity isn't mediated by competition for E-boxes (neither with Myc itself, nor with others TFs). Nonetheless, this mutant associated with open chromatin with profiles similar to those of their WT counterparts but does not position correctly on E-boxes. Yet, it drives aberrant transcriptional programs, which include not only the induction of new genes (see also our answer to Referee #3, point 9), but also a strong interference with E-box-containing Myc-activated genes.

The latter constitutes an apparent paradox, which it is particularly important to clarify here: indeed, as also pointed out by referee #3, the canonical CACGTG motif is significantly enriched among MycER^{HEA} downregulated genes (Fig. 5F), even though this mutant is not able to correctly recognize the E-box. We surmise that the enrichment of the CACGTG motif is not indicative of direct transcriptional repression, but rather of the fact that MycER^{HEA} is impeding the correct activity of endogenous Myc on those loci. This is consistent with the gene ontology analysis showing that MycER^{HEA} is repressing canonical Myc target genes (Fig. EV5C).

At the mechanistic level, the formation of Myc^{HEA}/Max dimers in cells is likely to have several consequences, which may all contribute to its DN activity. Moreover, these aberrant mechanisms may be important in the context of naturally occurring mutations in other bHLH subfamilies. As written in our Discussion (p. 13), "a series of mechanisms might contribute to the DN activity of Myc^{HEA}, such as (i.) titration of endogenous Max, (ii.) local hindrance of endogenous Myc/Max activity at promoters through non-specific DNA binding, (iii.) unproductive interaction with basal components or co-factors in the transcriptional machinery, or other mechanisms yet to be determined. Most intriguingly, mutations in the residue equivalent to Myc-E363 also conferred DN activity and altered DNA binding in other bHLH subclasses and species (from C. elegans to human) (Boisson et al., 2013, Luchtel et al., 2019, Marchegiani et al., 2015). Altogether, we surmise that the build-up of dysfunctional bHLH dimers may be detrimental within – and perhaps across – multiple bHLH families, warranting detailed dissection of their mechanisms of action."

As a final remark here, the aforementioned dominant-negative effects have no bearing to the main conclusions of our study, regarding the dual requirement of non-specific and specific DNA-binding for Myc activity.

Referee #1:

Using several solved structure of MYC, Pellanda et al generate two mutations in the basic region of murine Myc that are predicted to eliminate binding to DNA at either the phosphodiester backbone (MycRA) or the E box motif (MycHEA). The authors then asses the effects of these mutations by expressing the Myc mutants in different cell systems. The RA mutation impairs Myc's ability to bind chromatin, promote transcription, and drive cellular processes associated with Myc indicating that non-specific interaction with DNA is essential for Myc activity. The HEA mutation is more nuanced. When expressed at high levels it binds more genomic sites, increases transcription, and has a dominant-negative effect on WT Myc in cells. At endogenous levels MycHEA engages fewer genomic sites (especially with canonical E boxes), is crippled for cell proliferation, and prefers a novel E box derivative (CACGTC).

Based on their observations, the authors propose a model in which Myc binds to open chromatin in a sequence independent fashion followed by localization to E box elements that stabilize Myc and drive transcription. The authors' data sides with a model in which Myc activates specific genes rather than acting as a general transcriptional amplifier. The manuscript is interesting and the data would be of significant interest to the Myc community and likely to transcription regulation in general, but some issues need to be addressed.

Since the cellular levels of Myc correlate with binding to target genes, the manuscript could benefit from a more thorough characterization of the Myc mutations to insure that the only difference between the three versions of Myc is their affinity for E box motifs.

Indeed, the referee outlines a key point here: the biological effects of MYC in physiology and disease are determined largely by variations in its protein levels, which directly impact the binding to its target genes.

In this regard, we agree with the referee about the importance to make sure that the HEA and RA mutations impair the DNA binding, but not dimerization activities of MYC, as already suggested by our original co-IP data in transiently transfected 293T cells. We have now included several new experiments, which we believe thoroughly confirm this point. Please refer to our general answer for a description of those experiments (Authors' point 1).

Do these mutations alter the stability of Myc?

We addressed this with cycloheximide pulses followed by immunoblotting in two of our cell lines. The results are shown in Fig. EV2B and described in a new paragraph in the Results: as concluded there (p.5), "at this level of resolution, neither Myc^{HEA} nor Myc^{RA} showed altered protein stability"

Can the authors perform an assay that is more sensitive and quantitative than gel shift to determine the binding affinities to a canonical E box?

We have now refined our gel shift (EMSA) experiments with titration and competition assays, confirming the defect of Myc^{HEA} in E-box recognition (Fig. 1E-H). While other in vitro experiments may in principle be proposed here, we believe that the need for such experiments is largely circumvented by the extensive and quantitative characterization provided in our ChIP-seq analyses. Please see our general reply above (Authors' point 1) for further discussion of this aspect.

The authors show that MycRA and MycHEA engage Max at levels similar to the wild type protein (Fig EV1-C), but this is in an over-expression system. The authors should perform co-precipitation experiments for binding to Max ideally in cb9 cells, but at a minimum in the 3T9 system.

As mentioned above, this has been done.

Can the authors provide an explanation as to why MycHEA recognizes partial E-box motifs by ChIP (Fig 1J) but is unable to bind an Ebox by gel shift (EV1-D)?

This is a very relevant remark by the referee. As explained in our general reply (Authors' point 1), we have now repeated the EMSA experiments with independent proteins preparations and detected DNA binding by Myc^{HEA}/Max dimers.

If the author's suggestion that Max is responsible for recognition of an E-box half-site is correct, then why doesn't MycRA also recognize partial E box elements as both mutants still retain binding to Max?

This is in fact an important aspect of our data: stable binding to the E-box is determined by all of the atomic interactions between the proteins and the DNA, including backbone (non-specific) and base-specific interactions (Fig. 1B). Hence, it is logical to expect that loss of backbone interactions (as in Myc^{RA}) should abrogate all stable binding to DNA.

As also acknowledged by Referee #2 (point 1) and explained in our reply (see below), a mutant that loses non-specific binding but retains binding to the E-box would be very difficult – if not impossible to create.

Most noteworthy here, added the whole body of data presented, this is a feature that lends strong support to the sequential binding model put forward in our work: Myc^{RA} is unable to establish the first contact with DNA backbone, which obviously would also prevent sequence recognition (whether the half E-box or any other).

In Fig 2 changes in transcription are consistent with the presence of a canonical Ebox for MycWT or CACGTC for MycHEA in 3T9 cells. But do these transcriptional changes correlate with the distribution of MycHEA on chromatin from the 3T9 ChIPseq?

*This is a relevant question, and the answer is yes: as written in our text (p.10), "Myc-induced transcriptional programs correlated with the relative gain in Myc binding at promoters in diverse cell types (de Pretis et al., 2017, Lorenzin et al., 2016, Tesi et al., 2019, Walz et al., 2014). Likewise, up-regulated loci showed the strongest MycER-binding intensities in our experiments (**Fig. EV5A**), associated with enrichment of the cognate DNA motif (i. e. CACGTG for $MycER^{WT}$, CACGTC for $MycER^{HEA}$; **Fig. 5F, EV5B**)"*

An important concept made in our work is that the general distribution of ChIP-seq peaks correlates more extensively with general chromatin accessibility that with DNA sequence (p.6-7, Fig. 3A, Fig. EV3B,C). Yet, as written (p.6), "This notwithstanding, the motifs significantly contributed to the $MycER^{WT}$ profiles, as evidenced by three distinctive features: …" (please refer to the text for detail). As a corollary, transcriptional changes should not be merely correlated with the general distribution on chromatin, but with finer features, such as the quantitative gains in binding or the underlying sequence features.

Regarding the latter, in Fig. 5F (formerly Fig. 2F) we already showed that "the differences between $MycER^{WT}$- and $MycER^{HEA}$-driven transcriptional programs were attributable to DNA sequence…" (p.9). Most importantly, our work adds a key concept, which was not anticipated by former data, with the finding that, beyond binding strength per se, the presence of the cognate motif contributes to the transcriptional activity of the transcription factor on the corresponding promoter (Fig. 5H).

Additionally, the authors should perform qPCR on a selection of genes in cb9 cells and relate this to the cb9 ChIPseq results to determine whether the changes in transcription persist when MycHEA and MycRA are expressed at endogenous levels.

The cb9 cells are a valid model to address the genomic distribution of the retrovirally expressed Myc^{WT} and Myc^{HEA} proteins and their relationship to the underlying DNA sequence (Fig.EV3C-H), but are less suitable to dissect direct transcriptional responses. Indeed, these cells must first be grown with a large excess of Myc^{WT} expressed from the doxycycline-dependent tet-Myc transgene, which is then switched off. As a consequence, what one would be monitoring in this model are short-term changes in gene expression upon tet-Myc extinction, in the presence or absence of the retrovirally expressed proteins (Fig.EV1B,C): this is objectively a complex situation, with numerous confounding factors. Moreover, comparison of steady-state expression profiles at longer time-points would be equally inappropriate, as Myc^{HEA}-expressing cb9 cells arrest upon tet-Myc shutdown, while those with Myc^{WT} keep proliferating.

These considerations notwithstanding, we had already established RNA-seq profiles in Cb9 cells. In the scatter plots presented here (Fig. R3), we compare the differences in gene expression between cb9 cells expressing Myc^{WT} or Myc^{HEA} relative to control cells with the empty vector (EV), all profiled in the absence of doxycycline (i. e. with tet-Myc off). Strikingly, neither of these populations showed any correlation between changes in gene expression and DNA-binding intensity (left plots). Furthermore, neither the canonical E-box (CACGTG) nor the new motif (CACGTC) correlated with gene regulation in this model. Altogether, these data contrast with the clear correlations observed upon activation of the MycER proteins (Fig.EV5A,B), and confirm the inadequacy of the cb9 model to address changes in gene expression.

Given the lack of correlative data by RNA-seq, performing qPCR on a limited set of selected genes in the cb9 model – as suggested here by the referee – would not be representative, and would instead convey the risk of be being misleading. For these reasons, we decided not to present the cb9 expression data in our manuscript, but only the analysis of ChIP-seq profiles in these cells (Fig.EV3C-H). Most importantly, the latter were fully consistent with the analogous data in 3T9^{MycER} cells (Fig. 3A-D and I-J).

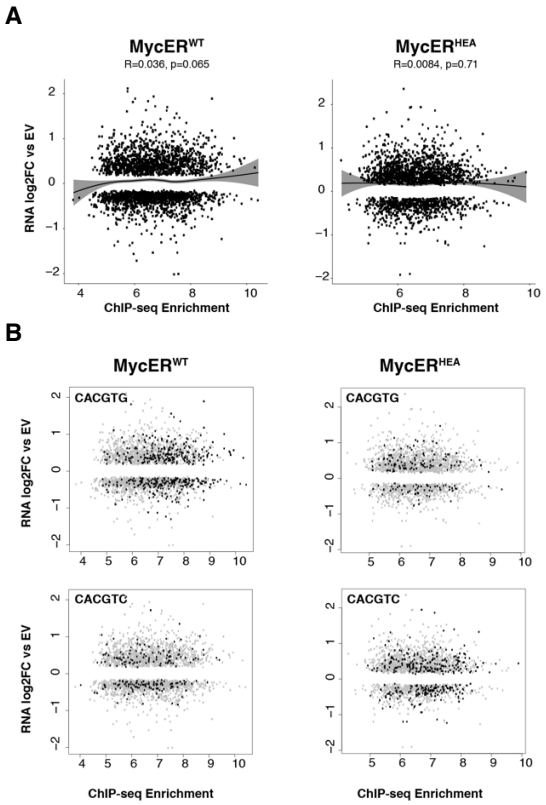

Figure R3: ***A.*** *scatter plots showing the fold-change of each mRNA (log2FC, Y-axis) against the enrichment of Myc binding to the corresponding promoter (X-axis) for all genes. The inclusion criteria for this comparison were (i.) for the mRNA to be called as differentially expressed (DEGs; qval<0.05) in cb9 cells expressing Myc^{WT} (left) or Myc^{HEA} (right) relative to control cells EV, all in the absence of doxycycline (16h), and (ii.) presence of a promoter-associated peak in the corresponding cells, as called by ChIP-seq.* ***B.*** *The plots show the same data, with black dots indicating the presence of either the canonical (CACGTG) or the new MycER^{HEA}-enriched motif (CACGTC) within ±100bp from the peak summit. Please note that these representations are analogous to those shown for MycER activation in 3T9 cells (Fig. EV5A, B).*

The immunoblot for Myc expression in cb9 cells without Dox (Fig EV3-B) is not convincing and should be performed again to show levels of MycWT.

The figure that the Referee mentions here contained both the blot that is now in Fig. EV1E, plus the original blot with the knock-in mutants that was removed from our paper (see above, Authors' point 3), Here, we must emphasize again that cb9$^{\Delta myc}$ cells express no endogenous Myc^{WT}, as the locus was inactivated by CRISPR/Cas9 engineering (Appendix Fig. S1A). Hence, the immunoblot presented in Fig. EV1B (originally Fig. EV3B) is perfectly adequate. Furthermore, in order to document Myc levels in the parental cb9 cells (i. e. before mutagenesis of the endogenous c-myc locus), the corresponding sample has been included in the new blot shown in Fig. EV1C).

On the other hand, as reported above (Authors' point 3), our original immunoblots assessing the levels of endogenously expressed Myc^{WT} and Myc^{HEA} were too preliminary, as subsequent experiments revealed defective expression of Myc^{HEA} in the knock-in cells (Fig. R2, above). This led us to remove the data with these clones from our manuscript. We thank the referee for bringing up this point.

How do the CACGTC sites bound by MycHEA in cb9 cells correlate with chromatin accessibility? Can the authors compare their cb9 ChIPseq to the published H3K4 and H3K27 data sets shown in Fig EV1-A? […]

The ChIP-seq profiles for the retrovirally expressed Myc proteins in cb9$^{\Delta myc}$ cells are now presented alongside the MycER, chromatin, RNAPII and DNaseI profiles in 3T9 cells (Fig. EV3C): this shows that,

like MycER^{HEA} (Fig. EV3B), Myc^{HEA} largely co-maps with its WT counterpart and with active/open chromatin.

Our detailed analyses further show that, as for the MycER forms (Fig. 3B-D and I-J), the binding of Myc^{WT} and Myc^{HEA} is fine-tuned by the CACGTG and CACGTC motifs, respectively: this is documented by motif frequencies (Fig. EV3D, G), peak intensities (Fig. EV3E) and positioning of the motifs within the peaks (Fig. EV3F, H).

[…] Is there anything else special about these sites?

No, there is nothing else special about those sites – and no formal need to postulate other special features.

It is interesting that the ability of Myc to bind chromatin is lost with the RA mutation even though MycRA should retain affinity for E boxes. Is it possible that non-specific contacts to DNA are required for Myc to stably bind DNA regardless of local scanning? In other words, minimal Myc binding at an E box requires the sum of non-specific and specific interactions with DNA.

Indeed, as already developed above, all or the atomic contacts between the proteins and the DNA are required for stable binding to the E-box: in this regard, contribution of non-specific backbone contacts isn't secondary to that of base-specific ones (Fig. EV1B).

Many transcription factors other than MYC recognize E box elements, yet MycHEA behaves as a dominant negative only in the presence of wild type Myc. Can the authors propose a mechanism for this level of specificity?

Please refer to our general answer (Authors' point 4) for our reply to this point.

Can the authors further elaborate on how their data supports a model of specific gene activation by Myc vs the general transcriptional amplifier model? This is an important conclusion given the controversy between these two models in the literature.

We thank the referee for bringing this up, and have now inserted a dedicated paragraph on this topic in our Discussion (p.12): "The specificity of MYC-dependent transcription became the subject of an active debate in the field…"

Referee #2:

The authors Pellanda & Dalsass et al. describe in their manuscript entitled "Integrated requirement of non-specific and sequence specific DNA binding in Myc-driven transcription" that non-specific binding of Myc to chromatin is a pre-requisite for subsequent sequence-specific E-box binding, and that sequence-specific binding is required for gene activation of specific transcriptional programs. The authors demonstrate that despite the ability of Myc mutants lacking E-box recognition to exhibit non-specific interactions with chromatin, that these interactions lack specificity and the ability to contribute to E-box associated gene activation and normal proliferative effects observed with wild-type Myc. Overall, the study is very interesting and sheds light on the clouded field of Myc DNA binding and transcriptional regulation.

The authors utilize Myc mutants unable to bind to DNA (MYCRA), and a MYCHEA mutant that lacks affinity for the canonical Myc E-box sites. These tools are powerful and informative. They also take advantage of cells engineered to delete endogenous Myc to allow for lower levels of expression of the mutants versus WT. At lower levels, the MYCHEA mutant has reduced affinity for DNA in general and some selection for a non-canonical E-box site. This results in a change in gene activation relative to WT Myc and the inability to support proliferation. These results support a role for sequence-specific binding in Myc transactivation activity.

Major concerns:

1) One of the authors' major claims, repeatedly stated throughout the manuscript, is that non-specific DNA binding is required for the subsequent sequence-specific binding to E-box sites. Even the title calls this an "integrated requirement". However, based on the two mutants used, it is not clear how this claim can be made since the mutants either lack DNA binding, or lack E-box specific binding. There is no mutant that can bind to E-box sites but not non-specifically to DNA (which would likely be difficult to create). This claim needs clarification to either keep the claim or to revise the interpretation.

The referee is right in noting that we are missing a mutant that would lack non-specific binding while still recognizing the E-box. However, while true from a theoretical perspective, this is simply not addressable at the experimental level. Indeed, having gone over this problem from the very beginning of our project, we deem it not be merely difficult (as acknowledged by the referee), but actually impossible to create such a mutant, since, as written in the initial section of our Results (p.4), "binding to the E-box should be supported by the sum of base- and backbone-directed interactions ***(Fig. 1B****): hence, while Myc^{HEA} might be predicted to retain non-specific binding, Myc^{RA} should lose all binding modalities: as shown below, both of these predictions were confirmed experimentally."*

This being said, the thorough characterization of the Myc^{HEA} and Myc^{RA} mutants with the combination of assays presented in our work makes a compelling case for the proposed model, and for the integrated requirement of the two binding modalities: please refer to our general answer above (Authors' point 2) for a detailed reply on this point.

2) If the claim is kept, mechanistic insights for how non-specific binding enhances the specific binding would be important.

As explained in our our general reply (Authors' point 2) and in answer to the previous point, the proposed model is strongly supported by the data, and is the only scenario that is fully consistent with all of the available results. On this basis, we deem is formally justified to maintain this model as the outcome of our paper. We have also kept a constant attention to producing all possible mechanistic understanding throughout our work, and believe that the data obtained at this stage make a compelling argument.

This being said, providing more "mechanistic insights for how non-specific binding enhances the specific binding", as requested here by the Referee, constitutes a whole new line of investigation: this would imply gaining a deep understanding of the biophysics of MYC-DNA interactions, addressing features such as on- and off-rates with different DNA motifs, mechanisms of lateral diffusion, etc… While we are planning new efforts in this direction, these are clearly beyond the scope of the present manuscript.

3) In figure 2B, the authors show 4 and 8 hr time points and differences in genes regulated by WT vs the HEA and RA Myc mutants. Can the authors identify a pattern of increasing gene expression over time with WT (or some temporal pattern) that is not apparent in the HEA mutant? This might shed light on transcriptional regulation mechanisms and gene programs lost with non-specific binding.

The heatmap shown in our original Fig. 2B (now Fig. 5B) showed the differences in genes expression, but was not suited to specifically visualize MycER^{WT}-regulated programs and their loss with the mutant proteins, since all of the samples together had been used to cluster the data. To illustrate this better, we now substituted this by two heatmaps, side-by-side: one in which the MycER^{WT} data drive the clustering (left), and a separate one driven by Myc^{HEA}. Most importantly, the differences between the MycER^{WT} and Myc^{HEA}-regulated programs are directly visualized in the dot plots presented in Fig. 5C-E.

As expected MycER^{WT} up- and down-regulated distinct sets of genes, and this response evolved over time, with numerous subgroups of genes showing heterogenous kinetics. As such, these profiles preclude the simple definition of defined "temporal patterns" as suggested by the referee.

4) There does appear to be some residual binding of MycHEA to E-box sites that the authors attribute to half sequence recognition via the heterodimer MAX. What about tetramers with endogenous MycWT? Is this still observed with the deletion of endogenous Myc? Likewise, the dominant negative activity of MycHEA with endogenous Myc present could reflect the HEA mutant reducing affinity of endogenous Myc in a tetramer with it from canonical sites needed for gene activation and proliferation. Further discussion would be useful.

Indeed, we had shown that MycER^{HEA} enriched for half E-boxes (Fig. 3F: previously 1J); of note, this was due mainly to distal binding sites, promoter proximal sites showing the complete variant consensus CACG(c/t)C (Fig. EV2D: previously Appendix Fig. S1B). The reviewer brought up the question of whether this is still observed with the deletion of endogenous Myc.

In Fig. EV3C, we showed the binding profiles for the retrovirally expressed MYC variants in cb9$^{\Delta myc}$ cells (that express no endogenous MYC) but had not presented de novo motif discovery on this dataset. This was due to the fact that the cb9$^{\Delta myc}$ cells provided no sizeable subset of Myc^{HEA}-only peaks – largely due to the lower level of "invasion" of active chromatin by Myc^{HEA} in this model, compared with MycER^{HEA} (Fig. EV3B-C).

As a corollary, unlike achieved with the MycER model, in which we could identify a sizeable subset of MycER^{HEA}-only peaks (Fig. EV3B) and run the motif discovery on those (Fig. 3H), this approach was not conclusive for Myc^{HEA}. While we did detect a relative enrichment of CAC half sites by de novo motif discovery on the overall Myc^{HEA} ChIP-seq (data not shown), as seen with MycER^{HEA} (Fig. 3F), this remained marginal and added no relevant information: we thus decided not to include this analysis in the paper.

Most importantly here, the need for de novo motif discovery in the in cb9$^{\Delta myc}$ cells was largely circumvented by our targeted analysis of the same dataset, which clearly showed that the binding of Myc^{WT} and Myc^{HEA} is fine-tuned by the CACGTG and CACGTC motifs, respectively (Fig. EV3). In particular, this is documented by motif frequencies (Fig. EV3D, G), peak intensities (Fig. EV3E) and positioning of the motifs within the peaks (Fig. EV3F, H), and essentially confirms what concluded in the 3T9^{MycER} model (Fig. 3B-D and I-J).

Finally, we fully agree with the referee with the potential interest of the hetero-tetrameric form of Myc/Max, described in the original crystal structure (Ferre-D'Amare et al., 1993). In fact, at the onset of our project we intended to use structure-based mutagenesis with two main goals (i.) to interfere with select DNA-binding steps and (ii.) to impair the putative tetramerization activity of Myc/Max dimers. While the first goal was pursued productively (as presented in this manuscript), the second wasn't – and was thus discontinued. In particular, examination of the structures did not allow us to design mutants that could be clearly be predicted to impair tetramerization. More critically even, the crystallographic evidence for Myc/Max heterotetramers (Ferre-D'Amare et al., 1993) was never corroborated by complementary experimental evidence, and remains a possible crystal-packing artifact. Because of these uncertainties, and given the consistency of our ChIP-seq data in 3T9^{MycER} and cb9$^{\Delta myc}$ models (with and without endogenous Myc, respectively), we feel that speculating on the possible role of tetramers wouldn't be formally grounded in our paper.

5) The data in EV5C showing that the MycHEA mutant can repress Myc regulated genes is striking and suggests that the MycHEA mutant antagonistically occupies these sites without transcriptional activation for some reason. But, in lines with the above comment, if this mutant does not specifically bind to the Ebox sites, how is it doing this? It would be interesting to know from the ChIP-seq data if the MycHEA mutant truly occupies these sites, which could help with mechanism and would support that Myc binding to specific DNA motifs is permissive to transcriptional activation, as opposed to general occupancy by Myc at the enhancer/promoter regions.

Please refer to our general answer (Authors' point 4) for our reply to this point.

6) In Figure EV2B the authors show greater invasion of chromatin with the MycERHEA mutant, and wish to determine whether this is due to the aberrant overexpression of the Myc mutants. In figure EV3 the authors utilize alteration of the endogenous alleles of cb9 cells with the HEA modification, and compare this to the enforced retroviral expression of the mutants with silenced endogenous expression of Myc. There is significantly less binding with the HEA mutant without endogenous Myc and even less with the HEA knock-in. Further discussion of this relative to what is shown in Fig 1 with greater general binding with the HEA mutant would be helpful as the prominence of Fig 1 leaves the reader expecting that the HEA mutant has more widespread binding.

As outlined in our general answer above (Authors' point 3) the ChIP-seq profiles with the Myc^{HEA} knock-in could not be considered as formally conclusive, and were removed from the paper.

On the other hand, the higher levels of "invasion" of active chromatin by MycER^{HEA} vs. MycER^{WT}, as well as the opposite patterns obtained with Myc^{WT} vs. Myc^{HEA}, remain key features of our data (Fig. EV3A-C). As described in out text (p.7),

*"this might be due to several experimental variables, including the slightly higher levels of the MycER vs. Myc proteins (**Fig. EV1C**) and their distinct activation modes (OHT-induced vs. steady-state)"*

7) The authors claim that sequence-specific DNA binding activates gene expression at three distinct levels: stabilization, positioning, and promotion of its transcriptional activity. It is not clear that these are distinct mechanisms, please clarify.

We have carefully reviewed these features in out text and figures, and feel that they are now fully clarified. Rather than reiterating it all here, we would kindly refer to our text, and in particular the section entitled "Sequence recognition determines transcriptional activation" (p.9-10).

8) If sequence-specific DNA binding stabilizes Myc to DNA then why does the WT and HEA both have equivalent decrease in mobility on DNA? Please discuss this further.

Our imaging data show that, while Myc^{RA} shows increased mobility, Myc^{HEA} does not, and is comparable to Myc^{WT} in this regard. Together with our ChIP-seq profiles, these data imply that the main feature that restrains diffusion in the nucleus is the non-specific binding to DNA at active chromatin domains, a feature that is retained in Myc^{HEA}, but not in Myc^{RA}. Please refer to our general reply (Authors' point 2) and to the new section on our Results ("Non-specific DNA binding restrains free diffusion of MYC in the nucleoplasm") for a full assessment of this aspect.

9) The HEA retroviral expression and HEA knock-in without endogenous Myc show substantially reduced general DNA binding activity (EV3C). Especially compared to the WT knock-in, which is expressed even lower and looks to be barely detectable in EV3B. Further discussion on this and the role of endogenous Myc in the HEA mutant's DNA binding in ChIPseq experiments with endogenous Myc and in the interpretation of the HEA data showing loss of Myc gene programs and proliferation is important.

As explained in our general reply (Authors' point 3), the data with the Myc^{HEA} knock-in were considered inconclusive and removed from the paper. Regarding the role of endogenous Myc, please refer to our detailed reply to the Referee's point no. 4, above.

10) Further, comparing proliferation between the MycWT and HEA with loss of endogenous Myc, which shows HEA reduced proliferation activity, please discuss whether the overall reduction in its DNA binding shown in EV3C could be an explanation.

As explained above, the Myc^{HEA} knock-in could no longer be used to support an "overall reduction" in DNA-binding activity, and was removed from the paper.

Our data show that, while still binding DNA non-specifically, Myc^{HEA} has lost the ability to selectively regulate Myc^{WT}-target genes: this fundamental defect in the activity of Myc^{HEA} is the most likely explanation of its loss of proliferative activity.

11) Page 8 line 10, down regulated genes by WT or HEA lacked E-box sites, but the next sentence says the HEA mutant represses Myc targets that have E-box sites. Please clarify.

We have now re-phrased this as follows in the last paragraph of the Results section (p10-11):

*"Finally, unlike activated genes, those down-regulated by either $MycER^{WT}$ or $MycER^{HEA}$ recruited the transcription factor with the lowest efficiency and lacked enrichment of the cognate binding motif (**Fig. EV5A, B**, **Fig. 5F**). Hence, as previously proposed (Baluapuri et al., 2019, de Pretis et al., 2017, Kaur & Cole, 2013), repression by either $MycER^{WT}$ or $MycER^{HEA}$ may be largely indirect. Of particular notice here, while $MycER^{HEA}$ did not bind the canonical CACGTG E-box, $MycER^{HEA}$-repressed genes enriched for this motif (**Fig. 5F**) as well as for known Myc-dependent gene signatures (**Fig. EV5C**), in line with the dominant-negative action of this mutant over endogenous Myc^{WT} (**Fig. 2A-C**, **Fig. EV1G**)."*

For further clarification of this aspect, please refer to our explanations on the dominant-negative activity of Myc^{HEA}, both in this rebuttal (Authors' point 4) and in our text.

Minor points:
1. It would be interesting to speculate as to why the HEA mutant is expressed at higher levels compared to WT in the cb9 knock-ins.

As clarified above (Authors' point 3) the Myc^{HEA} knock-in clones expressed in fact less, not more protein that the control Myc^{HEA} clones, and were discontinued.

2. On page 5, line 3, it is not clear what "binding hierarchy" is referring to, please clarify.

We replaced "binding hierarchy" by "peak intensity", which was the actual meaning (now in p.6).

3. On page 8 lines 3 and 5 reference figures S6A and S6B, although there are no such figures.

This has been corrected: Fig. EV5A and EV5B.

Referee #3:

This is a fascinating and important paper that addresses previously unexplored yet fundamental activities of the Myc protein. The authors set out to address the question of how non-specific and specific DNA-binding activities of Myc affect its genomic binding and transcriptional regulation. By mutating residues that contact either the DNA backbone (MYC RA) or the E-box consensus motif (MYC HEA), the authors show that non-specific DNA binding is required for Myc to bind to the genome, while specific sequence recognition is essential for maintaining Myc's normal transcriptional profiles and biological activity. These findings have relevance for interpreting ChIP_Seq data for MYC as well as for understanding how Myc influences expression of its target genes. The data presented are in general thorough and compelling although several issues should be addressed.

1. In Fig. EV1D, it would be important to show whether the Myc mutants bind probes non-canonical E-box probes in vitro?

Rather than using mutant probes, we have chosen to confront the activities of unlabeled oligonucleotides with different core motifs (CACGTG, CACGTC and GGATCC) in competing with the labeled CACGTG probe for binding to the various protein complexes (Fig, 1F-H). Please refer to our general answer to all referees (Authors' point 1) and our text (p.9) for a description of those experiments.

Why does MycHEA display some binding to E-box DNA in vivo but not in vitro?

We have now detected DNA binding by MycHEA also in vitro: please refer to our general answer above (Authors' point 1) for further detail on this point.

2. In Fig. 1A and EV2B, it appears that MycERHEA binds genomic DNA more strongly and widely in vivo than MycERWT? What is the explanation for this effect?

*Indeed, we report higher levels of "invasion" of active chromatin by $MycER^{HEA}$ vs. $MycER^{WT}$, but also an opposite pattern with Myc^{WT} vs. Myc^{HEA} (Fig. EV3B-C). Hence, these patterns do not reflect an intrinsic property of the HEA mutation, but rather the specific conditions encountered in each particular experiment. In this instance, as explained in out text (p.7), "this might be due to several experimental variables, including the slightly higher levels of the MycER vs. Myc proteins (**Fig. EV1C**) and their distinct activation modes (OHT-induced vs. steady-state)"*

Most importantly, it should be noted that this general association of Myc (or other TFs) with chromatin strongly correlates with global chromatin accessibility (as marked by DNAse I sensitivity), an active chromatin state (histone marks) and the presence of the basal transcription machinery (RNAPII), as also clearly illustrated by our data (Fig. EV3B, C).

3. In Fig. 1C, the differences in peak intensities among canonical E-box, variants and motif-free regions for MycHEA appear similar to MycWT. Does this weaken the conclusion that preferential binding of E-boxes is lost in MycHEA expressing cells.

*The reviewer is right in pointing this out. In this regard, it should be considered that relating ChIP-seq intensity to the underlying sequence is confounded by the fact ChIP-seq profiles represent the compound of multiple cross-links at the population level, including specific and non-specific binding events. Indeed, as reported in multiple studies and as also noted in our text (p.6), "peak intensity at MycER-binding sites correlated primarily with chromatin and RNAPII (**Fig. EV3B**), rather than with the presence of these consensus motifs (**Fig. 3A**)"*

Concerning our original Fig. 1C (now Fig. 3C), we reasoned that reducing the size of the DNA window from ±100 to ±50 bp from the peak summit for the scoring of ChIP-seq intensities should improve the specificity: indeed, as shown now in the figure, this yielded a more consistent difference between $MycER^{WT}$ and $MycER^{HEA}$.

Finally, note here that this is only one of the features that points to the loss of E-box recognition by $MycER^{HEA}$ in vivo, complementing the frequencies of the various motifs within peaks (Fig. 3B, I), their distances from the peak summit (Fig. 3D, J), and de novo motif analysis (Fig. 3E-H).

4. Fig. EV3B and 3C are a somewhat confusing. Is retroviral expression of Myc variants considered as overexpression like the MycER system? If so, why is there less binding with MycHEA expression compared to MycWT (even though MycHEA expression is higher), which is opposite from the MycER results. Interestingly, cb9-mycHEA is more highly expressed than cb9-mycWT, but still binds less. Is it

because MycHEA requires WT Myc to be able to better bind to DNA, or it needs to reach a certain threshold of expression to be able to bind DNA better than MycWT? It would be helpful to have Western blots comparing expression levels of Myc variants in all three systems (MycER, retroviral expression, and cb9-myc).

We agree with the importance of this point. The original Western blots were produced separately, and did not provide an accurate cross-comparison of the relative levels of the MycER and full-length Myc variants. Please refer to our new immunoblot in Fig. EV1C, and to our reply to point no. 2 above.

As mentioned in our general reply (Authors' point 3), the Myc^{HEA} knock-in clones were removed from our study.

The Referee pointed to a possible role of endogenous Myc in DNA binding by Myc^{HEA}: while this is an interesting concept, we deem that the available data do not provide a rational basis evoke this scenario. For a more detailed explanation, please refer to our answer to Referee #2 (point 4).

5. In Fig. 1F and EV4D, is the increase in residence time with HEA significant compared to WT? If so, why?

We have completed our plots with all statistical comparisons (now Fig. 4B, C). The apparent increase in residence time with HEA relative to WT (Fig. 4C) is not significant.

6. Related to the above points - although in several immunoblots the levels of the WT and mutant Myc are similar it would be reasonable for the authors to actually determine the half-lives of the WT and mutant proteins in the cells used here. Myc proteolysis has been suggested to be linked to its transcriptional activity. If the mutant proteins were stabilized then it could explain changes in protein mobility and possibly other aspects of the behavior of these proteins.

We addressed this with cycloheximide pulses followed by immunoblotting in two of our cell lines. The results are shown in EV2B and described in a new paragraph in the Results (p. 5): as written there, "at this level of resolution, neither Myc^{HEA} nor Myc^{RA} showed altered protein stability"

7. Is CACGTC bound motif more enriched in cb9-mycHEA cells than in MycERHEA cells? […]

The observed frequencies of the CACGTC motif are 10.5% with $MycER^{HEA}$ (Fig. 3I), and 15.1% in $cb9^{\Delta myc}$ cells expressing Myc^{HEA} (Fig. EV3G). The differences between those numbers may be due to the higher levels of general "invasion" seen with the MycER form, reflecting a higher proportion of non-specific binding events. However, considering that these ChIP-seq data were generated with different antibodies (anti-ER and anti-Myc, respectively), we would formally restrain from inferring too much from this comparison. Most importantly here, the common feature shared between $MycER^{HEA}$ and Myc^{HEA} is that both selectively enrich for the alternative CACGTC motif, as opposed to the E-boxes (CACGTG and variants #2-5).

[…] Is the enrichment for degenerate AC-rich motifs lost in the cb9-mycHEA cells?

We cannot provide a formal answer on this point, as de novo motif analysis on the $cb9^{\Delta myc}$ cells expressing Myc^{HEA} was inconclusive: for a more detailed explanation, please refer to our answer to Referee #2 (point 4)

8. In Fig. 2F and EV5B, is CACGTG motif significantly enriched in down-regulated genes in MycERHEA cells? If so it would be consistent with the dominant negative effect of MycHEA.

As the referee correctly pointed out, the CACGTG motif is significantly enriched in the $MycER^{HEA}$ downregulated genes (we have now added the corresponding pval to Fig 5F - was 2F) consistent with its dominant negative role. For full clarification of this aspect, please refer to our general answer on the dominant-negative activity of Myc^{HEA} (Authors' point 4).

9. What specific growth related and metabolic pathways are positively enriched in MycHEA cells? Does MycHEA inhibit cell proliferation mainly by failure to activate Myc target genes or by activation of a different subset of growth-suppressive genes?

As discussed in our general reply (Authors' point 4), the dominant-negative activity of Myc^{HEA} is most likely attributable to its negative impact on Myc-activated genes.

On the other hand, regarding possible pathways positively enriched by MycER^{HEA} , gene ontology analysis showed that among the HALLMARK dataset, the following were significantly enriched (FDR<0.05):

HALLMARK_HYPOXIA
HALLMARK_P53_PATHWAY
HALLMARK_FATTY_ACID_METABOLISM
HALLMARK_ADIPOGENESIS
HALLMARK_XENOBIOTIC_METABOLISM
HALLMARK_ESTROGEN_RESPONSE_EARLY
HALLMARK_HEME_METABOLISM
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_APOPTOSIS
HALLMARK_ESTROGEN_RESPONSE_LATE
HALLMARK_GLYCOLYSIS
HALLMARK_MYOGENESIS
HALLMARK_TNFA_SIGNALING_VIA_NFKB
HALLMARK_SPERMATOGENESIS
HALLMARK_WNT_BETA_CATENIN_SIGNALING
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_PEROXISOME
HALLMARK_BILE_ACID_METABOLISM
HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_PROTEIN_SECRETION

Objectively, in the absence of additional evidence, developing a discussion on the role of any of these pathways appears premature, and potentially misleading. For example, considering that the 3T9 fibroblasts used in our experiment do not have wild-type p53, we cannot simply ascribe the growth inhibitory effects of MycER^{HEA} to the category "p53 pathway". Other categories such as "apoptosis", "hypoxia", "interferon gamma response" or "interferon alpha response" might be pertinent in this context, but would remain purely speculative.

Most importantly, none of the above impinge on the key concepts made in our work. As such, the characterization of the aberrant biological activities acquired by the MycER^{HEA} mutant – or, as mentioned in our Discussion, equivalent DN mutants in other bHLH families – is beyond the scope of our paper.

Minor:

1. In Fig. EV1C, is the shift of MycRA significant compared to MycWT?

This is now Fig. EV1F: this shift of Myc^{RA} was not consistently observed in other blots, and in particular not on the very same samples following IP (see the lower blot in the same figure). We thus deem it an artifact of this particular gel.

2. On page 8, line 5-8, Fig. S6A or Fig. S6B are not included in the manuscript.

This has been corrected: Fig. EV5A and EV5B

3. What's the point of showing the "spike in" in EV1?

The referee is right in saying that there was little point in showing the "spike in" version of the ChIP-seq profiles obtained from 3T9-MycER cells with Myc-specific antibodies (now in Fig. EV3A): these immunoprecipitates are ambiguous by definition, as they include signals stemming from the endogenous Myc and exogenous MycER proteins, without any possibility to discriminate between the two.

Precisely for this reason, the anti-Myc ChIP-seq profiles were not used to make any specific point in our paper. Instead, these samples served as controls, highlighting their close similarity with the profiles obtained with the anti-ER antibody, while the latter allows selective detection of exogenous MycER (Fig. EV3A: "anti-Myc" vs. "anti-ER"). We have thus removed the lanes with the spike-in analysis of anti-Myc immunoprecipitates.

Thank you for submitting your revised manuscript, we have now received the reports from the initial referees (see comments below). I am pleased to say that the referees find that their comments have been satisfactorily addressed and now support publication. I would like to therefore ask you to address a number of editorial issues that are listed in detail below. Please make any changes to the manuscript text in the attached document only using the "track changes" option. Once these remaining issues are resolved, we will be happy to formally accept the manuscript for publication.

Referee #1:

The revised manuscript by Pellenda and colleagues has undergone extensive technical enhancement as well significant and valuable changes to the text. The main conclusions of the manuscript are thoroughly supported by the data. The work makes a compelling case for the idea that both sequence specific and non-specific DNA contacts are required for target gene recognition by Myc, and that these two modes have different functional outcomes. The work will be of wide interest to the Myc community.

Referee #2:

the authors have done an excellent job of addressing all of my concerns and i believe the paper strongly supports the innovative findings and i recommend publication.

Referee #3:

The revised manuscript is significantly improved and the authors have responded in a satisfactory manner to all my comments. I consider the paper to be a valuable contribution to the MYC field.

2nd Revision - Editorial Decision 24th Feb 2021

Thank you again for submitting the final revised version of your manuscript. I am pleased to inform you that we have now accepted it 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: Bruno Amati and Arianna Sabò
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2020-105464

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

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 iustified
- 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.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

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?	As specified in Materials and Methods: "All experiments were performed at least in biological triplicates. Sample size was not predetermined, but is reported in the respective Figure legends."
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Not applicable
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Not applicable
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.	Not applicable
For animal studies, include a statement about randomization even if no randomization was used.	Not applicable
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.	Not applicable
4.b. For animal studies, include a statement about blinding even if no blinding was done	Not applicable
5. For every figure, are statistical tests justified as appropriate?	For every figure the appropriate statistical test is used as indicated in the legend.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	As specified in Materials and Methods: "Wilcoxon's test is used for non-normal distributions. Fisher's exact test is used for categorical data. Chi-squared test is used for categorical data where the contingency table is 2xN where N > 2. In the ramaining cases, two tailed-Student t-test was used to compare between two groups and expressed as p-values. Each test was used when data met their assumptions of validity."
Is there an estimate of variation within each group of data?	Not performed

USEFUL LINKS FOR COMPLETING THIS FORM

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

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-tumo

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/

Is the variance similar between the groups that are being statistically compared?	Not performed

C- Reagents

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).	All the antibodies used have been described in the table "Antibodies" in Matherial and Methods
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	As specified in Materials and Methods: "All cell lines were routinely tested for mycoplasma contamination"

* 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.	Not applicable
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	Not applicable
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.	Not applicable

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	Not applicable
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.	Not applicable
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	Not applicable
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	Not applicable
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	Not applicable
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.	Not applicable
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.	Not applicable

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for '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	See the "Data Availability" section in our article
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 the "Data Availability" section in our article
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).	Not applicable
21. 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.	Code not deposited

G- Dual use research of concern

22. 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.	Not applicable