

Eduardo Gade Gusmão <eggduzao@gmail.com>

BIOINF-2013-1895.R1 - Minor Revision

bioinformatics.editorialoffice@oup.com
bioinformatics.editorialoffice@oup.com> 29 de maio de 2014 02:01 Para: eduardo.gusmao@rwth-aachen.de

Cc: eduardo.gusmao@rwth-aachen.de, Christoph.Dieterich@age.mpg.de, martin.zenke@rwth-aachen.de, ivan.costa@rwth-aachen.de, igcf@cin.ufpe.br

28-May-2014

Manuscript ID: BIOINF-2013-1895.R1

Title: Detection of Active Transcription Factor Binding Sites with the Combination of DNase Hypersensitivity and **Histone Modifications**

Dear Mr. Gusmão.

The reviews of your manuscript are now in hand. On behalf of the Bioinformatics Associate Editor, Inanc Birol, I am pleased to report that the reviewers found your manuscript acceptable for publication in Bioinformatics, subject to the suggested revisions. I am therefore returning your original manuscript for revision in line with the editorial and reviewer comments which can be found at the foot of this e-mail.

Please submit your revised version through the Author Center by clicking on the purple button 'Click here to Submit a Revision' in the Bioinformatics ScholarOne Manuscripts web site (http://mc.manuscriptcentral. com/bioinformatics).

Please ensure that you use either mandatory template format which can be found at:http://www.oxfordjournals.org/bioinformatics/for authors/submission online.html

We ask that revisions are submitted within one month. The system will automatically remove the revision option if a revised paper has not been submitted within this time.

To facilitate the production process we ask that you upload the following revised manuscript files at the revision stage:

EITHER: (i) A .doc or .rtf file of the revised manuscript, with all tables, figures, schemes and equations inserted in the document.

OR: (ii) All necessary LaTeX files that will be required by the typesetter (including bioinfo.cls, bib, .bst and .ps files) along with postscript and PDF versions of the complete manuscript.

Please can you mark-up the changes made after revision by using the track changes function or highlighting these in red text.

Please upload your final clean version of supplementary materials with your revised submission. This should be in pdf or Word format, not LaTex.

I would also ask that you prepare an accompanying letter explaining exactly how each of the major points raised by the reviewers was addressed. This can be done either in a file uploaded alongside your revised manuscript as a Response to Reviewers file or through the Author Center where you can enter your responses directly in the appropriate box during the revision submission.

Please note that if you decide that you would like your figures printed in colour a charge of £350 per colour figure applies. If appropriate, you will be invoiced after your paper has been published in the print journal.

As a reminder, please also note the following excess page charges. You will be notified of any excess page charges when you receive your proofs:

For Original articles - £100/\$165 per excess page (over 7 published pages)

For Discovery notes - £100/\$165 per excess page (over 4 published pages)

For Application notes - £100/\$165 per excess page (over 2 published pages)

On behalf of the Executive Editor, I want to thank you for selecting Bioinformatics to present your work.

Best regards. Alison Hutchins **Bioinformatics**

Here are the comments of the reviewers:

Reviewer: 1

Comments to the Author

The paper is much improved and represents a useful piece of work. However I would still request the following changes before acceptance.

a)major

- 1) I would be much happier if all references to "ROC curve" were changed to "ROC-like curve" in the text and captions of the main and supplemental text. Even though you have now noted in one place that you don't use the standard definition, using the term "ROC" is misleading and confusing. Fortunately "AUC" works out mathematically to be the same for both ROC curves and your "ROC-like" curves, so that needn't be changed. Definitions are important.
- 2) To talk about specificity or sensitivity of a method you need to fix one of them. Please clarify in Table 1 how you picked a single value of specificity (sensitivity) for each method when applying the Friedmann-Nemenyi test. Also, in the text, please explain what it means to say that "Cuellar presented higher sensitivity and lower specificity". I'm not sure that is even a meaningful statement. At all specificity levels in Fig 2, Cuellar seems to have lower sensitivity. So state in what cases (and in what sense) Cuellar gives higher sensitivity.

b) minor

Typos:

p 10 line 5 "enhancer elements"

p 10 line 40 "open (and closed)"

p 11 line 27 "reads to have 200 bp" --> reads to be 200 bp long

Reviewer: 2

Comments to the Author

The authors have addressed most of the comments from the previous review, but there are a few points remaining:

- 1. The manuscript makes a fundamental assumption about active binding sites with regard to histone modification state around the binding sites. It is not clear what fraction of TFs follow this pattern. Are there specific families where this observation is not necessarily true?
- 2. On a related note, what is the false negative rate of the method? The training is based on manual annotation of a particular locus and therefore does not capture the binding sites that do not follow the histone modification pattern. One way to look at this is to analyze the DNase hotspots not annotated to contain footprints by the HMM method. Are there are specific motifs enriched in these peaks?

3. The results also show multiple footprints within a single peak -- do each of these footprints correspond to the binding of a different factor or is this an artifact of the way the HMM is set up? This is important information for downstream analyses such as motif discovery.

Reviewer: 3

Comments to the Author

The authors have addressed my concerns. I would ask for a couple of clarifications -while the footprints can be detected without knowledge of the TF, their interpretation does obviously require a known PWM. This is not clear in all cases.

Also, from a biological point of view, it is highly surprising that ES cells would not need separate parameterization (but maybe the relevant parameters are in fact rolled into the preprocessing of the data...). Its chromatin landscape is very different from differentiated cells (cf de Wit, Bouwman et al Nature 2013), and it might be worth to add a comment in this regard.

Ivan Gesteira Costa Filho <ivan.costa@rwth-aachen.de>

29 de maio de 2014 06:29

Para: Eduardo Gade Gusmão <eggduzao@gmail.com>

parabens!

depois dou uma olhada com calma nos comentarios!

[Texto das mensagens anteriores oculto]

Ivan G. Costa

www.costalab.org

IZKF Computational Biology Research Group RWTH University Hospital Aachen

- > doctoral stipends in computational engineering available in
- > graduate school AICES see http://www.aices.eu

Martin Zenke <martin.zenke@rwth-aachen.de>

29 de maio de 2014 18:57

Para: Eduardo Gade Gusmão <eduardo.gusmao@rwth-aachen.de>, Ivan Gesteira Costa Filho <ivan.costa@rwthaachen.de>

Dear Eduardo and dear Ivan,

having read the reviewers' further questions/concerns I think it is doable. Please let me know, whether I can be of further help, in particular for the points of Reviewers 2 and 3.

Best regards

Martin

Begin forwarded message:

From: <bioinformatics.editorialoffice@oup.com> Date: May 29, 2014 2:01:28 AM GMT+02:00 To: <eduardo.gusmao@rwth-aachen.de>

Cc: <eduardo.gusmao@rwth-aachen.de>, <Christoph.Dieterich@age.mpg.de>,

<martin.zenke@rwth-aachen.de>, <ivan.costa@rwth-aachen.de>,

<igcf@cin.ufpe.br>

Subject: BIOINF-2013-1895.R1 - Minor Revision

[Texto das mensagens anteriores oculto]

Martin Zenke, PhD, Professor of Cell Biology and Chairman Institute for Biomedical Engineering Department of Cell Biology Universitätsklinikum Aachen, RWTH Pauwelsstrasse 30 52074 Aachen

Helmholtz Institute for Biomedical Engineering Pauwelsstrasse 20 52074 Aachen Germany

Voice +49-241-80 80759 (secretary) +49-241-80 80760 (office)

Fax +49-241-80 82008

martin.zenke@rwth-aachen.de email: andrea.offergeld@rwth-aachen.de

http://www.molcell.de www:

Christoph Dieterich MPI-AGE < Christoph. Dieterich@age.mpg.de> 30 de maio de 2014 07:36 Para: Eduardo Gade Gusmão <eduardo.gusmao@rwth-aachen.de>, Ivan Gesteira Costa Filho <ivan.costa@rwthaachen.de>

Great Job!

We made it.

All best wishes,

Christoph

P.S.: Thanks Ivan for inviting me towards the Brazil workshop.

Dr. Christoph Dieterich

Max-Planck-Institut für Biologie des Alterns Max Planck Institute for Biology of Ageing

Leiter Bioinformatik / **Head of Bioinformatics**

Office: Joseph-Stelzmann Straße 9b,

D-50931 Köln / Cologne

Postal address: Postfach 41 06 23, D-50866 Köln / Cologne

Tel.: +49 0 221 37970 310

christoph.dieterich@age.mpg.de www.age.mpg.de

Am 29.05.2014 um 02:01 schrieb bioinformatics.editorialoffice@oup.com:

[Texto das mensagens anteriores oculto]

Ivan Gesteira Costa Filho <ivan.costa@rwth-aachen.de> Para: Eduardo Gade Gusmão <eggduzao@gmail.com>

30 de maio de 2014 08:29

oi,

acho que o paper ja praticamente passou. Revisores 1 e 3 estao obviamente ok, so o 2 precisa de alguma argumentacao e algumas analises adicionais nos resultados. Podemos tambem por mais resultados para a questao de ES do revisor 3.

> reviwer 2

Este ta guereno encrecar um pouco, mas acho que nao vai ser problema, ja que os outros 2 gostaram ...

>1. The manuscript makes a fundamental assumption about active binding sites with regard to histone modification state around the binding sites. It is not clear what >fraction of TFs follow this pattern. Are there specific families where this observation is not necessarily true?

O que poderiamos fazer en pegar a classificação de TF do

http://nar.oxfordjournals.org/content/41/D1/D165.long

eh associar com o AUC individuais de cada fator. Outra coisa que podemos analizar eh a associacao do AUC com o info. content.

>2. On a related note, what is the false negative rate of the method? The training is based on manual annotation of a particular locus and therefore does not capture >the binding sites that do not follow the histone modification pattern. One way to look at this is to analyze the DNase hotspots not annotated to contain footprints by >the HMM method. Are there are specific motifs enriched in these peaks?

Isso eh valido (e possivelmente simples de fazer). De outro lado o FPR pode ser calculado na nossa metodologia de avaliacao.

>3. The results also show multiple footprints within a single peak -- do each of these footprints correspond to the binding of a different factor or is this an artifact of the >way the HMM is set up? This is important information for downstream analyses such as motif discovery.

Obivamente acreditamos no cobinding, mas isto eh dificil de argumentar ia que nao existe maneira experimental de provar cobinding. Poderiamos avaliar quanto porcento dos footrings tem TP para alguns dos fatores, mas obivamente so far sentido se tivermos medido todos os possiveis TFs da celula. Talves para k562 valha a pena. Outra coisa que podemos fazer eh comparar o numero de footrints em promoter/enhancer e nos CTCFs (que deve ser sozinhos).

> Reviewer: 3

>Also, from a biological point of view, it is highly surprising that ES cells would not need separate parameterization > (but maybe the relevant parameters are in fact rolled into the preprocessing of the data...). Its chromatin >landscape is very different from differentiated cells (cf de Wit, Bouwman et al Nature 2013), and it might be worth > to add a comment in this regard.

Podemos reviver aquela analise das histonas separadas por celula.

Ufa, quase lah. Entao, qual a data mesmo da sua volta? Nao estou vendo as datas no calendario do grupo!

abs,

Ivan

[Texto das mensagens anteriores oculto] [Texto das mensagens anteriores oculto]

Ivan Gesteira Costa Filho <ivan.costa@rwth-aachen.de> Para: Christoph Dieterich MPI-AGE < Christoph. Dieterich@age.mpg.de> Cc: Eduardo Gade Gusmão <eduardo.gusmao@rwth-aachen.de>

30 de maio de 2014 08:43

Hi,

Yes, it is looking good and finger crossed for the brazilian workshop.

best,

Ivan

2014-05-30 7:36 GMT+02:00 Christoph Dieterich MPI-AGE

<Christoph.Dieterich@age.mpg.de>:

[Texto das mensagens anteriores oculto]

[Texto das mensagens anteriores oculto]