Author's response to reviews

Title: Unraveling genomic variation from next generation sequencing data

Authors:

Georgios A Pavlopoulos (g.pavlopoulos@med.uoc.gr)
Anastasis Oulas (oulas@imbb.forth.gr)
Ernesto Iacucci (ernesto.iacucci@gmail.com)
Alejandro Sifrim (alejandro.sifrim@esat.kuleuven.de)
Yves Moreau (yves.moreau@esat.kuleuven.de)
Reinhard Schneider (reinhard.schneider@uni.lu)
Jan Aerts (jan.aerts@esat.kuleuven.de)
Ioannis Iliopoulos (iliopj@med.uoc.gr)

Version: 3 Date: 10 June 2013

Author's response to reviews: see over
Dear editors,

We herewith would like to submit the revised manuscript which is now entitled:

“Unraveling genomic variation from next generation sequencing data”

Below you can find a point-by-point reply with all of the suggested revisions addressed.

We would like to thank you in advance for your attention and for reviewing the current study, and we would be grateful if you could consider this manuscript for publication in your journal.

Looking forward to your decision,

Yours sincerely,

The authors
1. The abstract does not reflect the title. Either the abstract has to be re-written to suit the title and the main text or vice versa. Alternatively, one possible suggestion for an appropriate title, based on the abstract and the main text, is “Unraveling genomic variation from next generation sequencing data”.

We thank the reviewer for the comment. The title has now changed accordingly. Indeed the new title is shorter and more to the point.

Previous title: Unraveling the structural variome using insights from next generation sequencing technology: tools and tips for a messy job

New title: Unraveling genomic variation from next generation sequencing data

2. Several citations are incorrect in the paragraph under Sequencing Techniques. For example, a review on gene networks is cited for SOLiD technology. The authors should cite the primary articles instead of obscure papers that just mention the technology. References 22-31 need to be changed and primary or most appropriate review (for example, cite a Church or Shendure paper for second generation sequencing) cited. I don’t think Sanger sequencing is next generation.

We thank the reviewer for pointing out these errors. The references are now updated and pointing to original publications. Indeed Sanger technique is not next-generation and this is now fixed and clarified within the text.

3. Next generation sequencing applications: The paragraph title is out of place since the previous title is way generalized. Definition of “next generation” should be made carefully in the previous section.

We thank the reviewer for this comment. As it can be shown also in comment 5 the section “Sequencing techniques” and “Next generation sequencing applications” are now merged into one as they are complementary to each other and cannot be really differentiated. Therefore, the structure of the document has changed. Now the top header is: “Sequencing technologies” which is divided in two smaller sections. The first one is entitled: “1st, 2nd and 3rd generation” and “DNA sequencing and assembly”

4. ChIP-seq information is shortchanged. I think it would be good to remove RNAseq and ChIP-seq sections. The references are mostly wrong, again, please check all the references (reference 50 is wrong, Reference 57 is not correct).

…part of comment 4 transferred here: RNA sequencing (better to eliminate RNAseq - make a mention of other extensions by citing a review (for example, take one from Mike Snyder)

We thank the reviewer for giving us a clearer image of how the content of the manuscript should change in order to be more understandable. The ChIP-seq and RNA-seq sections are now completely removed and the references got updated. We mention these possible application shortly and we redirect the readers to extensive reviews for both RNA-seq (Mike Snyder) and ChIP-seq (PMID: 23090257).
The headings, somehow, do not make sense. I wonder if the following headings would suit the text, in the order it is presented.

1. Introduction
2. Sequencing technologies
   a. Applications of next generation sequencing technologies
   b. DNA sequencing
      i. Types of sequence assemblies
      ii. File formats
3. Genome annotation and variant calling
   a. SNP calling pipelines
   b. Structural variation detection
4. Discussion

We thank a lot the reviewer for his suggestions. Taking a second look at this comment we realized that the titles of each section should be clearer and more descriptive. Therefore the current structure of the manuscript is as follows (see previous and new structure with changes highlighted):

Current Scheme:

1. Introduction
2. Sequencing technologies
   a. 1st, 2nd and 3rd generation
   b. DNA sequencing and assembly
3. The structural variome
4. Methods to detect structural variations
   a. Pair-end mapping
   b. Single-end
   c. Split-reads
5. File formats
   a. FASTQ
   b. Sequence Alignment/Map (SAM) format
   c. Variant Call Format (VCF)
6. Variant calling pipelines
7. Variant annotation
8. Visualization of structural variation
   a. Alignment tools
   b. Genome browsers
   c. Visualization for comparative genomics
9. Discussion

Previous Scheme:

1. Introduction
2. Sequencing techniques
3. Next generation sequencing applications
   a. DNA sequencing and assembly
   b. RNA sequencing
      i. Filtering
      ii. Assembly
      iii. Expression profiling
      iv. Annotation
   c. ChIP-sequencing
4. SNPs, CNVs and the structural variome
5. Discovering structural variations using next generation sequencing
   a. Pair-end mapping
   b. Single-end
   c. Split-reads
6. File formats
   a. FASTQ
   b. Sequence Alignment/Map (SAM) format
   c. Variant Call Format (VCF)
7. Variant callers
8. Variant annotators
9. Visualization tools
   a. Alignment tools
   b. Genome browsers
   c. Visualization for comparative genomics
10. Discussion

6. Additional changes done by us

   • Some parts of the text that were not adding any important information were removed. These parts are highlighted in red.

   • Typos and smaller corrections were made and highlighted in red.

   • Figure 5 and its description made simpler.

   • Few of the tools that seem to be outdated were removed from the tables.

   • Many of the references were wrong. That was a big mistake from our side as two different EndNote libraries were confused. We thank the reviewer for pointing out this issue. Now the references are added one by one from scratch and updated.