Peer Review File

Article information: https://dx.doi.org/aob-21-39

Reviewer A

In this paper, the author mainly introduces the feature and types of next generation sequencing (NGS) and their applications in blood group genotyping.

The first part "introduction" is about the importance of NGS in transfusion medicine. NGS has many applications and can circumvent limitations of currently routinely-used detection methods like serological tests.

Then, the author briefly introduces the development of BGA sequencing, NGS platforms applied for BGA genotyping (sequencing-by-synthesis offered by Illumina and semiconductor sequencing on Ion Torrent from Thermo fisher) and how the two NGS platforms works.

The core content of this paper is "BGA NGS strategies". In this section, the author summarizes current studies of the applications of four strategies including targeted NGS, whole-genome sequencing (WGS), whole-exome sequencing (WES) and third generation sequencing in blood group genotyping. Also, the author compares different methods and run down their pros and cons, providing guidance for future usage.

"Genetic background of BGAs and NGS" gives us a simple view of genetic changes responsible for antigen variety. Three types are mentioned. These are SNVs-determined, enzyme-related polysaccharide antigens like ABO system and deletions-based, for example, RHD system.

Finally, the author concludes the applications of BGA NGS, and offer a summary of current research on blood type by NGS.

Highlights

This paper gives a great introduction for BGA NGS strategies. It comprehensively summarizes current studies of the applications of the four strategies in blood group genotyping field. As a review, it provides us with complete information on the

progress of research in this field. Besides, it also compares different methods and tell the advantages and limitations of each of them, offering guidance for future usage.

Limitations

Comment 1:

There are a few points in this paper that could be better. Firstly, for the section "Development of BGA sequencing", the content doesn't match the title. The text is primarily about the feature and flexibility of BGA sequencing and is also a bit related with the application.

Reply 1: I agree that the section was a mix of the different content that didn't match the title. I deleted the title as in my opinion the paragraph is a part of the introduction section.

Changes in the text: page 4 line 84 "Development of BGA sequencing" deleted

Comment 2:

Next, in "BGA NGS strategies", some WGS-based works are placed in the WES part.

Reply 2: The cited works were checked and corrected.

Changes in the text: page 10 line 241-243 moved to WGS-based section: page 9 line 215-217

Comment 3:

Last but not the least, although it covers all kinds of genetic changes responsible for antigen variety, the "genetic background of BGAs and NGS" is a little bit simple.

Reply 3: The paragraph is only a brief view on the background of BGAs as it is not a main topic thus it is rather short just to give an impression on diverse molecular structures that are sequenced and analyzed. Not to repeat some limitations associated with the BGA genetics in line 277 page 7 I put a statement "…as mentioned above" in the previous sections. But I added some details as the Reviewer suggested.

Changes in the text: page 12 line 275, page 12 line 280

Reviewer B

The present manuscripts reviews the next-generation-sequencing-based genotyping technologies and the strategies used for blood group antigen typing and discusses their advantages, the possibilities these technologies have opened up, and the challenges that come with using them. The review comprehensively summarizes the relevant literature, lays out the current state of the field and points towards future directions.

The manuscript is generally well written, although may benefit from proofreading by a native English speaker, layed out clearly, and a timely contribution to the field of immunohaematology.

Minor points of improvement:

Comment 1:

Page 5: if one points out that the per-read error rate on Illumina systems is around 0.1%, one should also point out that the per-read error-rate on Ion Torrent is generally substantially higher (>=1%), albeit the errors are mostly indel errors. It would also be prudent to include information on the capacity (read-length, output) for a comparable Illumina system (e.g., the MiSeq).

Reply 1: All comments were taken into account and included.

Changes in the text: page lines 112-113 added on Illumina systems, line 121 on Ion Torrent

Comment 2:

Page 10: not to be nitpicky, but the language in the section on third-gen sequencing technologies (PacBio and ONT) should be revised for clarity and grammatical/factual correctness. Several phrases, whilst broadly unterstandable, are used in a quite unusual manner (e.g., "omits the limits of short reads", "we meet two technical solutions", "offers sequencing")

Also "electric changes" => "change in electric current", "passes through a membrane

with protein nanopores" => "passes through protein nanopores embedded in a membrane"

With ONT sequencing the electric current does not change with each nucleotide passing through the pore as put in the manuscript, but rather with every 5- or 6-mer that occupies the pore at any given point in earlier pore versions, and three central nucleotides plus smaller influences from more distal nucleotides at later pore versions (since R9, R9.4). (I'm not suggesting that you put all these details in the manuscript but try to be more exact)

Also maybe it should be pointed out in this section that, at least to my knowledge, nobody has yet attempted or at least published about using PacBio or ONT sequencing for bloodgroup typing

Reply 2: The language was revised and the incorrect phrases were changed according to the Reviewer's suggestions.

Changes in the text: page 10 lines 244-256

Comment 3:

Page 13: "analogous to HLA"

Reply 3: changed according to the Reviewer's suggestions.

Changes in the text: page 12 line 292

Reviewer C

This is a well-presented review of the current status of NGS in blood group genotyping. My comments on the manuscript are relatively trivial and wholly designed to improve quality, which is already of a high standard.

(1) Page 8 Line 18, 22, 25: The three "BG" abbreviations used in this paragraph mean different, I recommend it should be modified, ex: (Line 18) 36 "BGs" should mean 36 "blood group system" according to reference 44; (Line 22) 45 "BG" should mean 45

“blood group related genes” according to reference 46; (Line 25) 38 “BG” should mean 38 “blood group antigens” according to reference 47.

Reply 1: All abbreviations were changed according to the Reviewer’s suggestions.

Changes in the text: page 9 lines 199-206

(2) Page 11 Line 12: "e.g. ABO*01" should be O allele (ABO*O.01)

Reply 2: changed according to the Reviewer’s suggestions.

Changes in the text: page 11 line 272

(3) Page 23 reference 78: The year of this reference should be 2017.

Reply 3: changed according to the Reviewer’s suggestions.

Changes in the text: page 24 line 569