ROUND 2

Reviewer 1

Basic reporting

I have no further questions.

Experimental design

The authors present a tool which loads, annotates and prioritizes variants loaded using standard VCF format. Experiments are basically analysis of available data sets and comparing with what is known in the literature, which are appropriate for this study.

Validity of the findings

As described above, findings are compared with the literature

Comments for the author

Minor: Galen et al paper is now published online in Bioinformatics. Please update references.

Our response

Dear Reviewer 1,
We have updated the reference accordingly, many thanks.

Reviewer 2

Basic reporting

The manuscript is clearly written and is easy to understand and follow.

Experimental design

The authors described their prioritization approach in the methods. However, the manuscript lacks information on the gene fusion part of NGB.

E.g. how do you detect the gene fusions? Is it a greedy approach to say any SV connecting two genes? Do you allow for temporary hops to other regions?

Validity of the findings

All fine.

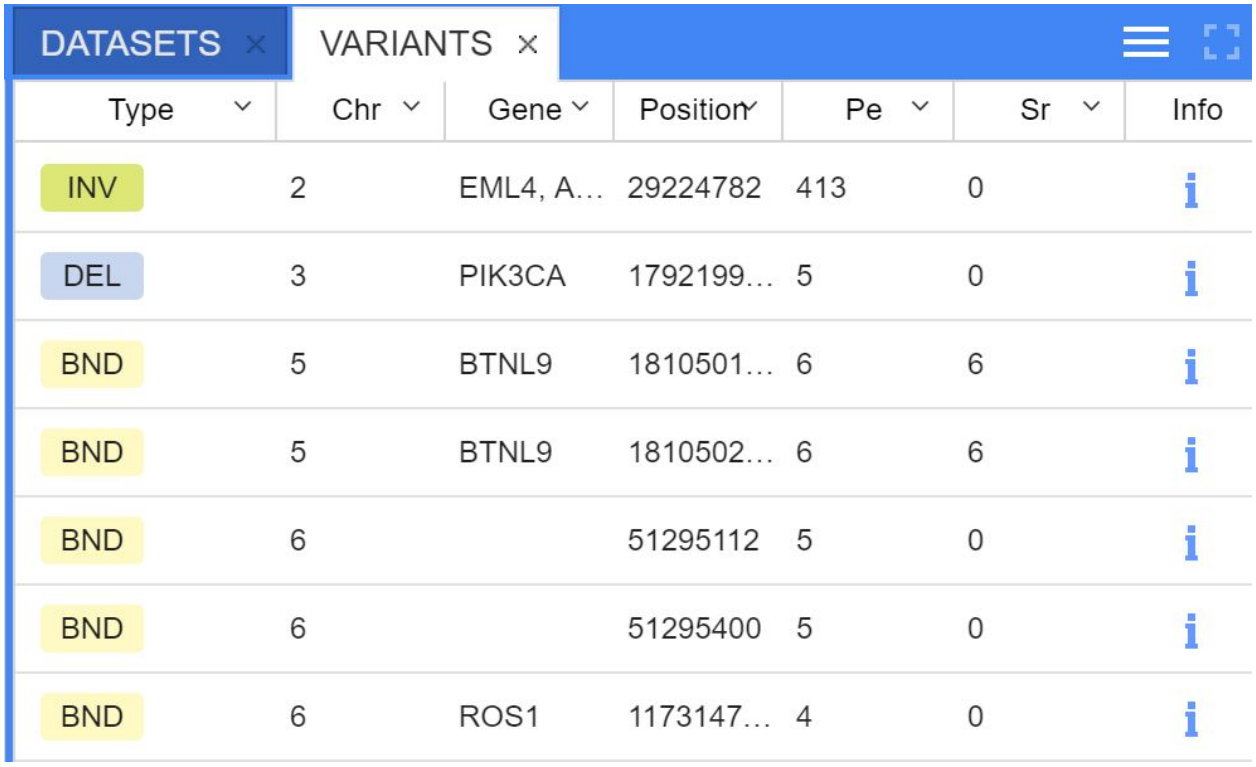
Comments for the author

I really like the NGB gene fusion visualization. I would suggest that you also report the number of supported reads in the visualization. Since you already require this information it would help users to understand if this gene fusion is well supported or not.

Our response

Dear Reviewer 2,
Just to clarify, NGB does not perform any fusion prediction or detection but merely reads information in the vcf file (mainly the SnpEff annotation). The text “We ... implement a variant call based gene fusion visualisation scheme in the open source New Genome Browser ... NGB takes the variant breakpoints **annotated by SnpEff** and uses Ensembl and UniProt based annotation to visualise the fusion product in both reference as well as the actual sequence context” ought to reflect this but if it is not clear enough please let us know. At the moment only simple events are considered where the breakpoints affect two genes (and are considered a fusion by SnpEff). Without long reads and phasing it would be very difficult to predict the effects of separate SVs affecting the same allele of a gene (i.e. hopping). We welcome suggestions on how to practically facilitate this on https://github.com/epam/NGB

Regarding the number of supported reads in the visualisation, there is already support for showing all vcf fields in the variants list and filter/sort by them. I am attaching a screenshot here showing “PE” (paired end reads) and “SR” (split reads), which were both selected from the “three horizontal bars” pop up menu top right:



Type	Chr	Gene	Position	Pe	Sr	Info
INV	2	EML4, A…	29224782	413	0	i
DEL	3	PIK3CA	1792199…	5	0	i
BND	5	BTNL9	1810501…	6	6	i
BND	5	BTNL9	1810502…	6	6	i
BND	6		51295112	5	0	i
BND	6		51295400	5	0	i
BND	6	ROS1	1173147…	4	0	i

We have recently modified the SV-simplification program to provide an INFO field of the highest priority of all possible effects of a single variant to help sort the variants better in NGB. For any further queries please open tickets on Github.