From: Wen Yao <ywhzau@gmail.com>
Time: 2014/1/19 11:42
To: 'babak aref'
Subject: **Re**: Re: Re: Re: Re: Re: Re: Re: Re: Assist

Dear Babak,
You are welcome.
Thanks for your suggestion. I will add that feature to intansv in the next release.

Regards,
Wen

From: babak aref [mailto:]
Time: 2014/1/19 11:10
To: Wen Yao
Subject: Re: Re: Re: Re: Re: Re: Re: Re: Re: Assist

Dear Wen,
Thank you for your explanation. Yes, I saw abnormal read pairs in that regions. I wanna design primers for checking the SV experimentally. So, I must pick up homozygotes. Thanks.
For bettere prediction I filter out regions for genomic gaps and telomeric regions (for my investigation horse genome has not defined centromere). If you think providing this option for intansv would be possible, It will make it super package for structural variation survey.

Once again Thank you for your time.

Regards,
Babak

From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Sunday, January 19, 2014 6:21 AM
Subject: Re: Re: Re: Re: Re: Re: Re: Re: Assist

Dear Babak,
Pvalue was not used in the filtering of CNVnator's output since different experiment gives different pvalue ranges. I'm not sure how to choose the proper threshold. And several published papers used CNVnator didn't use pvalue to filter.
As for the coverage for deletion regions, this might be the results of repetitive sequences or it might be heterozygous deletion. I think you could check the coverage of deletion regions and compared it with whole genome average coverage or the

coverage of normal regions. You can also check the abnormal read pairs that support the presence of a deletion. If a deletion was supported by a cluster of abnormal read pairs, I prefer it as a deletion even there were read coverages within it.

Regards,
Wen

From: babak aref []
Time: 2014/1/19 10:14
To: Wen Yao
Subject: Re: Re: Re: Re: Re: Re: Re: Re: Assist

Dear Wen,
Thank you.
For cnvnator does it has any filtering option for pvalue?
When I checked the deletion result from sv_merged_all_methods with genomeview (for coverage plot of tdf generated from bam), I saw that most of the deletions have coverage within the related regions. Do you have any comment for this?

Regards,
Babak

From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Sunday, January 19, 2014 5:19 AM
Subject: Re: Re: Re: Re: Re: Re: Re: Assist

Dear Babak,

The attachment file contains two figures with legends for your question. I hope it’ s helpful to you.

Regards,
Wen

From: babak aref [mailto:]
Time: 2014/1/18 2:51
To: Wen Yao
Subject: Re: Re: Re: Re: Re: Re: Re: Assist

Dear Wen,
Would you please tell me that how intansv filter the cnvnator, pindel and

breakdancer data? Which criteria are implemented during filtering and merging them?

Regards,
Babak

From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Friday, January 10, 2014 1:49 PM
Subject: Re: Re: Re: Re: Re: Re: Assist

Dear Babak,
You are welcome. If you encounter any problems using intansv in the future, please contact me.

Regards,
Wen

From: babak aref [mailto:]
Time: 2014/1/10 16:53
To: Wen Yao
Subject: Re: Re: Re: Re: Re: Re: Assist

Dear Wen,
Thank you for your time. It sounds nice. I appreciate your assistance.

Regards,
Babak
From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Thursday, January 9, 2014 5:34 PM
Subject: Re: Re: Re: Re: Re: Assist

Dear Babak,

I have fixed the bugs in intansv. Now you can do the annotation and plotChromosome. However your "genome" (genome <- GRangesForUCSCGenome...) is not quite proper, I had modified it to genome.1. The attachment is the modified package and the results of your data as RData and pdf file.

Regards,
Wen

From: babak aref [mailto:]
Time: 2014/1/6 12:00
To: Wen Yao
Subject: Re: Re: Re: Re: Re: Assist

Many thanks

From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Monday, January 6, 2014 7:21 AM
Subject: Re: Re: Re: Re: Assist

I' ll fix this as soon as possible.

From: babak aref [mailto:]
Time: 2014/1/5 22:27
To: Wen Yao
Subject: Re: Re: Re: Re: Assist

Dear Wen,
Thanks. It works.
I have created gff and genome GRange objects by this:

```
>EquCab2_71 <- import.gff("EquCab2_71.gff3")
>genome <- GRangesForUCSCGenome("equCab2",chrom=c("chr1", "chr2", "chr3",
"chr4", "chr5", "chr6", "chr7", "chr8", "chr9", "chr10", "chr11", "chr12", "chr13",
"chr14", "chr15", "chr16", "chr17", "chr18", "chr19", "chr20", "chr21", "chr22",
"chr23", "chr24", "chr25", "chr26", "chr27", "chr28", "chr29", "chr30", "chr31",
"chrX"))
```

I have attached my gff and also RData files.
When I tried to annotate sv_all_methods with my gff by this:

```
sv_all_methods.anno <- llply
(sv_all_methods,svAnnotation,genomeAnnotation=EquCab2_71)
```

I encountered this error:

```
Error in validObject(.Object) :
  invalid class "GRanges" object: 'seqnames' contains missing values
```

and when I tried to plotChromosome I had this error.

```
> plotChromosome(genome,sv_all_methods,1000000)
Error in dupDf$chromosome : $ operator is invalid for atomic vectors
```

Regards,
Babak

From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Sunday, January 5, 2014 4:54 PM
Subject: Re: Re: Re: Assist

Dear Babak,

Is intansv okay now?

Regards

Wen Yao

From: babak aref [mailto:]
Time: 2014/1/5 19:27
To: Wen Yao
Subject: Re: Re: Re: Assist

Dear Wen,
Thank you.
I will check right now.

Regards,
Babak

From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Sunday, January 5, 2014 7:39 AM
Subject: Re: Re: Assist

Dear Babak,

I have fixed the bugs in readDelly and methodsMerge. I have update intansv in Bioconductor but it will take a few hours to make the change come into effect. So I send to you the source and binary package of the updated intansv in the attachment. Now you can read your delly output into R. And you can merge the output of as few as one method now.

If you encounter any problems or have any suggestions, please contact me.

Best regards.
Wen Yao

From: babak aref [mailto:]
Time: 2014/1/5/ 10:25
To: Wen Yao
Subject: Re: Re: Assist

Dear Wen,
Thank you for your response.
With the example data I found that all the methods are required. So, I tried to perform delly and svseq. Unfortunately, the problem duplicated and the delly output couldn't import using readDelly function. It encountered this error during the import: "Error in dellyData[, 1:7] : incorrect number of dimensions". I have attached my delly files. I could not find the source of the error in my data. Its format is as like as example data.

Regards,
Babak

From: Wen Yao <ywhzau@gmail.com>
To: 'babak aref' <>
Sent: Sunday, January 5, 2014 5:16 AM
Subject: Re: Assist

Hi Babak,

Thanks for your interest in intansv. I’ m sorry for bring you this trouble. This is a bug and I’ ll fix it as soon as possible. I’ ll notice you then.

Regards.

Wen Yao

From: babak aref [mailto:]
Time: 2014/1/4/ 3:42

To: ywhzau@gmail.com
Subject: Assist

Dear Wen,
Happy new year.
I have tried to perform "intansv" R package for my breakdancer and cnvnator results. When I perform the below script I encountered the error in non matching name length. Therefore, I would like to ask you if you could please assist me to solve this problem.

Your assistance is greatly appreciated.

```
sv_all_methods <- methodsMerge
(breakdancer=breakdancer,pindel=NULL,cnvnator=cnvnator,delly=NULL,svseq=NULL)

Error in names(InversionDfMerge)[1:3] <- c("chromosome", "pos1", "pos2") :
  'names' attribute [3] must be the same length as the vector [0]
```

Regards,
Babak