agrep deletion bug testing

10-Jan-2024

```
version info:
$ agrep -V
AGREP 3.41.5/TG for NATIVE LINUX compiled with GCC 10.3.0 (Nov 8 2023 18:20:55).
Manber/Wu/Gries et al.
$ cat /etc/os-release
NAME="CentOS Linux"
VERSION="7 (Core)"
ID="centos"
ID_LIKE="rhel fedora"
...
```

When using agrep on DNA sequence files (FASTQ format), some records output by agrep had deletions of varying length, with each deletion occurring precisely after the first 1023 characters (or 1021 characters plus a newline, in the examples below).

To reproduce the problem, I created a file agrep_testc_30.txt with 30 copies of a record (varying only the index numbers); this is a record suffered a deletion in real data.

Here is the prototypical record: (sequence preceding deletions highlighted in blue) sample id=Pan 21s002174 A TAAGATGTAAGACATGAATGCACATAAGATTATTTATAGTAACTAAACATATTGGATTAGTCACTGCACGGAATGATGTTAGCAATGTCTTCCTGGTTTTGAAATCTTAAATGAACTACAAGGAAGCGCTGACTCGGCTATCCAACTTTCGGCCATCAAGCAGGACAAGCCTAACCATCGAAACCCAGGAACAATAACCATCAAGGCAGGAACCATCCAAC $\texttt{GCCCATCAAAAGTGTAGCCAATAAAATCAGAGGATCGTAGAGAAGGTGATCCAGCTACAGAGAGTGATATTGATAATTGATCATAAATTGATCAACAAATTAT$ %%),.1/--))+*+'&%(\$%(#\$(&%%&54969<>>;>983)7==>AHAOI:@DA6...,,,,*****)))(((((('%/+7<B6,''&&&&&&'%)-23-336696..*59179:=>D@CFG=CAD@>;;9&\$)00+),-8;;557<9:84<>1+60?=9>E0?<11()&,.;A8=<75)1%&&20/02+(.%.09:;7D7DJDEHF<==5'%&)5=3,244:6;>0151,/0027;:9><(''*/&%&&&))()*)*)8##8/00(7.)*.5667-32))688:CAD@:(78,7,4754(''),((1***./)\$\$'4016.,64=95<;;79>*+**.0+82:76.''78:>==4312/\$'(28*)/97;88=:<;<)(*-,'5,44((,/.+6++,)(./+1::@B@9?5ACEDF@<C54.5=9<:8>?CB4CGGKHB9-''(4:@HE;6.,--(-,1113:0??A?C>F;BG950??009/%#'.1322/)*01.8;>>?=67>:0FCB>B:;55*6;?AEE??B804?*'2.54%494503-0111.(4)%&&&*2463*-48<9+',%)(\$%+%(:.\$\$1\$%379:53*++'\$,-.''(',2:8)8?@98<>5,6;60/-&%%.3774++,-/*))+'(*-/13539@=673598;<C;A?=;42341.7264469:>HDA>C@7;;?;7;;8:561)0:;6<=>88:6%*/-)&&8+41??=3657:/1('\$+%'%&%,691%+*(/78:? @@AB@A9.618+\$%')/681-,*3501134++*+,%+*'\$(''')(\$\$'#\$#\$&''('('%&&'\$\$'%%%%\$\$\$\$\$\$\$\$\$\$\$\$\$\$\$\##%#\$\$&\)()\$\$'\$%+)%.'*'%%%\$)+, -,+'\$\$##'(**)(%%\$##%\$\$#&\$''(('&'\$#\$\$\$\$##\$\$\$\$&&&'\\$\$##\)-,'(//.)44/0,(,,-,,*%##.//2.'0242',/1-1,-31/22,\$#&*+%),-*\$&340(&0/.,+(%#&&'\$%(#(%%\$\$&&&'&%%\$\$%%\$

Here is a command that reproduces the problem. The agrep command matches on every record and outputs every record, but three of the output records (10, 19, 29) have deletions; they are given below.

```
$ rcstr="TTAGATGGTTCCTGGGGTTCTGATGGC"
$ agrep -2 -d "^@" $rcstr agrep_test_30.txt >agrep_test_30.out
```

```
TACCGTGTTCCAATGTCAAAGTTTTTTTTTTTTTTGACGGAGGGCGAAAAGTGAAGTGTACATGG
'&(265>FE?<-0%**-'$55AAHHE=A6HLI<<=6173@=?;=:*8AA373,49664301./..-..../././...-.----
00111055.998.0&%%'2&45E@>HDGF583-
/,&$$+%884'06<<;8:A8D><<ELO685A?@MJ<@B?AHF6=7901=>=???9<?JG?IN<CDEGFHD7<HADGKKD9AKDLF4,4.))+675;<7,/037>>@<<?7&'
&2*-(''0469;88,=@CFCA3&)9:?DBB@==3B34883:::3)),=><@><<6454,&'&$.-
78:?:9F<B89:;B,+347@A<=**;JJGGA?=<0%2256,0&73/)))(('''''$03:8%%&(223/2779:6<::;77;4>C@=5424560144<75+-
2477@<7;:C?><CKAA?@DCEI<HC;:>9<:67A=<EC(...77;34.3;;>>><;4651?IUD@F@;GDD?6-
2GG<45534CGGLIKLD@=:?EGCFGEC>?CCDC=:A;9;><=???;9<@A<FHD<GJ?KBBBG?G@PC:;817>?>B;EDC@:976,/@CF92DHF;=A>:@?&-
D?7;05;:86=C=9E@<=<77;??=??ADD:<DCF@o>?@;2'9><)79?@;<>7:6B=;:CB<30A=@?B?6,</c&&&):9?<?1:<B8>?B*?6:;2-
45.>)6>>><=BG;<)?8689===?>0*#&$,,-/*$$&&&,7-
41;:83<>E=CC::82)'))3789865'&$#%%(%)''(587:<;>C=89A6;F:AD>D?>?<C?CE=BB+/>;=?>IK.++'+@FG?17?979:<5@7B?-:E>C>$1.77
68?=A@FDHD9@>A9@9CA?:.C<=@<;;@;A=>@GI@A@?EAF?=?6@H33/2?=@1:<C4CB@HIEKJHHB=%(%$(=<;148823@CG85*303457@767446699@;
;7=798>9=<>@39<87563023&2430+)'%&'((.14::../%.6&:721.+=@=EI?BEDG668DD>C;;@8?89.26267==;971589:=???6</,532:DDCA94
50.42/\$\$\$-5/1.)-68B71(\$^++,-\$\$^+'^*\$^+'^*+\$^*\$^*))35\&)(35:787:(+.00463/62=5.\$;631'7101..)),-
-) &29=48>45>74.**37+(9-./23.0%, &2//2-
0211/+)) \\ \\ \$ \& \$ \& ``) \\ \\ // \& `('**., +(('\&)') \\ \\ \$ \& \& ``\$ \\ \\ \$ \% \\ \\ *`§ & ('*§ \$ \$ \$ * * \$ `) \\ \\ +/05' \\ \$ \& `' ` \$ \& \& ``§ ``` ` ()' ` \$ \$ * * * \& \& () , / `` \& () , / `` & )' \\ \\ \& ((*\$ \$ \$ * . * \& (*), * ) \\ \\ \& (*\$ \$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . * \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * ) \\ \\ \& (*\$ * . \& (*), * 
**))) **&) ').'*''$$$%&)&&*$$'%&(&%$$$&&++*--+.-.(''($/.-,$#"$%&$$&',$%&'))*'*../-0./4$...,:5+)41(+,..*$-//1(-
%'8AA:2+%,-.-)',$'&'&*0/1)&$&'('%%$$$$&)%(,271+0$'$'$(&**'&%$,/+1+('$#$#$'')+&$%(%#$%*(,+&'&()+-,,*))###$(./0:9-
6+&$$*()(,./59<483%'031----**)'(&$$$$$$#'(')()*/-552*)$$$#"$%$%$%&&&&$"&.+*%)&(((**&(,'&$#&*&(-332,
start_time=2021-05-31T19:53:39Z flow_cell_id=FAO40255 protocol_group_id=210531_FAO40255_Pan_21s002174
sample id=Pan 21s002174
\tt GTTCCTCTGGTTCTGATGGCTGGTTAGATGGTTCCGCAGGTTCTGGTGGCTGGTTAGATGGTTCCTGGAGGTTCTGATGGCAGCTGGATTTTGGATGGTTCCTGTTCTGATGGTTCTGATGGCAGCTGGATTTTGGATGGTTCCTCTGGTTCTGATGGTTCTGATGGCAGCTTGGATGGTTCCTGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTCTGATGGTTGATGGTTGATGGTTGATGGTTCTGATGGTTTGATGGTTGATGGTTGATGGTTGATGGTTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGGTTGATGG
TGGCTGGTC./0:9-6+&%$*()(,./59<483%'031----**)'(&$$$$$#'(')()*/-
start time=2021-05-31T19:53:39Z flow cell id=FAO40255 protocol group id=210531 FAO40255 Pan 21s002174
sample id=Pan 21s002174
\verb|CGTGTTCCAATGTCAAAGTTTTTTTTTTTTTTGACGGAGGGCGAAAAGTGAAGTGTACATGG|\\
```

I made test files from two other records, also with 30 copies. They both are records that suffered a deletion in real data.

```
This file agrep testb 30.txt didn't give any deletions.
start time=2021-05-31T19:18:03Z flow cell id=FAO40255 protocol group id=210531 FAO40255 Pan 21s002174
sample id=Pan 21s002174
\tt CTGGTTAGATGGTTCCTGGAGGTTCTGATGGCTGGTTAGATGGTTCCTGAAGTTTCTGATGGCTGGTTGGATGGGTTCTCTGATGGCTGGTCAGCTGGCTTCTTT
{\tt GGTTCAGGCTGATTCTTCATGTCAGAAGGATCCTTTTCTACATAGTTATGTCACAGCTCCCCCAGAATGGTACCTTACTTCCTGTGAATATTTCTGCAAGCCACAGCCATTT}
\tt CGCAGCTGCCGACAATCGTCCTACCATCTTGGCGATGATTAGAAGATTTAAAGGTTCAGCTACAATGGCTGTGGCTTCATGCAATGTTGTTGTAGTACCATTCTGGGAGTTG
\tt CTGGTGATGATAACTGTATATAGAAAAGGATCCTTCTGACACGAAGAATCAGCCTGAACAGGAAACCAGCTGACCAGCCATCAGAACCAGGAACCATCCAACCAGCCAT
&$%**):?>@@BA+&*&&***/,*,%:368:;*3:76@;;=E8@:C9@/;:9;;996621//..----%;//347;AAB=9$5:8@CD;9F>)9==5@ABC@<<7-
***$(##$'&&%(.3;32./((($%&&&'%&%#$%$%*5350/>=91$$.*--125:9?A@<FCFB?I:;D@><CA+?@@AB?;;::9>>@9:=,)14-4'269:/-
0;3321469A<54?>=;<>=6BA0:803(&4=;.::9:656/('-04:@A>>?F86.-))33<&2&('*024--
1=6466::4%49667%$%(,),-,,.80?=c?=5;109;?AGF;67-,-1+/..301*'$988=;2<990,-0096?B7545532,-
22.).:5,%$/*%(/%'/&0/..5445;:3008'%%&/25.)'&$&&,5445;70-0/286;-
79532''***') &+6212388 < (@AA == 7 << 7:56 + .7175 \\ \$ \\ \$ \\ (((\$ +) .2. - \&(\&(*, -, -)) +, , '. / (+, +2587 - :?5/9:7, \&('\&65/ - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. - .2. -
```

Command:

\$ agrep -2 -d "^@" \$rcstr \$SCRATCH/agrep testb 30.txt >\$SCRATCH/agrep testb 30.out

```
The third test file agrep testc 30.txt gave two deletions (records 12 and 23).
start time=2021-06-01T01:33:27Z flow cell id=FAO40255 protocol group id=210531 FAO40255 Pan 21s002174
sample id=Pan 21s002174
ATAAGATGTAAGACATGAATGCACATAAGATTATTATATAGTAACTAAACATATTGGATTAGTCACTGCACGGAATGATGTTAGCAATGTCTTCCTGGTTTTGAAATCTTAAA
{\tt TGGTTAGATGGTTCCTGGGGTTCTGATGGCTGGTTCCTCTGGTTCTGATGGCTGGTCGTGACACAGGTTTTCCAGGCTGATTCTTCGTGTCAGAAGGATCAGCC}
{\tt TGAACTACAAGGAAGCGCTGACTCGGCTATCCAACTTTCGGCCATCAAGCAGGACAACCCTAACCATCGAAACCCAGGAACCATCAAGGCAGGAACCATCCAACCTTCCAACCTACCATCCAACCTACCATCGAAACCCATCAAGGCAGGAACCATCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCCAACCTTCAACCTTCAACCTTCCAACCTTCCAACCTTCAACCTTCAACCTTCAACCTTCCAACCTTCAACCTTCAACCTTCAAC
\tt GCCCATCAAAAGTGTAGCCAATAAAATCAGAGGATCGTAGAGAAGGTGATCCAGCTACAGAGAGTGATATTGATAATTGATCATATATTGATCAACAAATTAT
%%),.1/--))+*+'&%($%($%($%(4$(&%%&54969<>>;>983)7==>AHAOI:@DA6...,,,,*****)))(((((('%/+7<B6,''&&&&&\%)-23-
336696..*59179:=>D@CFG=CAD@>;;9&$)00+),-
8;;557<9:84<>1+60?=9>E0?<11()&,.;A8=<75)1%&&20/02+(.%.09:;7D7DJDEHF<==5'%&)5=3,244:6;>0151,/0027;:9><(''*/&%&&&)
)()*)*)%##%/00(7.)*.5667-
32))688:CAD@:(78,7,4754(''),((1***./)$$'4016.,64=95<;;79>*+**.0+82:76.''78:>==4312/$'(28*)/97;88=:<;<)(*-,'5,44(
(,/.+6++,)(./+1::@B@9?5ACEDF@<C54.5=9<:8>?CB4CGGKHB9-''(4:@HE;6.,--
(-,1113:0??A?C>F;BG950??@09/\$\#".1322/)*01.8;>>?=67>:0FCB>B:;55*6;?AEE??B804?*"2.54\$494503-0111.(4)\\ %\&\&\&*2463*-1113:0?A?C>F;BG950??@09/\$\#".1322/)*01.8;>>?=67>:0FCB>B:;55*6;AEE?B804?*"2.54$494503-0111.(4)\\ %\&\&\&*2463*-1113:0?A?C>F;BG950??@09/8\#".1322/)*01.8;>>=67>:0FCB>B:;55*6;AEE?B804?*"2.54$494503-0111.(4)\\ %\&\&\&*2463*-1113:0?A?C>F;BG950??@09/8#".1322/)*01.8;>>=67>:0FCB>B:;55*6;AEE?B804?*"2.54$494503-0111.(4)\\ %\&\&\&*2463*-1113:0?A?C>F;BG950??@09/8#".1322/)*01.8;>>=67>:0FCB>B:;55*6;AEE?B804?*"2.54$494503-0111.(4)\\ %\&\&\&*2463*-1113:0?A?C>F;BG950?**0113:0.54$494503-0111.(4)\\ %\&\&\&*2463*-1113:0.54$494503-0111.(4)\\ %\&\&\&*2463*-1113:0.54$494503-0111.(4)\\ %\&\&\&A*CA*-1113:0.54$494503-0111.(4)\\ %\&\&\&A*CA*-1113:0.54$494503-0111.(4)\\ %\&\&\&A*CA*-1113:0.54$494503-0111.(4)\\ %\&\&A*CA*-1113:0.54$494503-0111.(4)\\ %\&\&A*-1113:0.54$494503-0111.(4)\\ %\&A*-1113:0.54$494503-0111.(4)\\ %\&A*-1113:0.54$49403-0111.(4)\\ %A*-1113:0.54$49403-0111.(4)\\ %A*-1113:0.54$49403-0111.(4)\\ %A*-1113:0.54$49403-01
48<9+',%) ($%+%(:.$$1$%379:53*++'$,-.''(',2:8)8?@98<>5,6;60/-&%%.3774++,-/*))+'(*-
/13539@=673598;<C;A?=;42341.7264469:>HDA>C@7;;?;7;;8:561)0:;6<=>88:6%*/-)&&8+41??=3657:/1('$+%'%&%%,691%+*(/78:?
@@AB@A9.618+$%')/681-,*3501134++*+,%+*'$(''')($$'#$#$&''('('%&%'$$'%%%$$$$$$$$$$$$$\##%#$$$&\)()$$'$%+)%.'*'%%%$)+,
&% (00.#%% (%$$$#&'&''&)--/1))/.,++''(%$%%$$$$$$$))*'%$&(#+))('#%%)(.(-
##%&&$#%&$$''$%%#&&&''''$$$%%%$$(&$$%('#$%\'\%#$\%\'\$$'\$\#\%\\\%\##$#\#$#\#$#\#$\#\$\$\&\\))##$$&\$)$%$&&\(&&%%%%%%$$\'
*$&340(&0/.,+(%#&&'$%(#(%%$$&&&'&%%$$%%$
Command:
$ rcstr="TTAGATGGTTCCTGGGGTTCTGATGGC"
$ agrep -2 -d "^@" $rcstr agrep testc 30.txt >agrep_testc_30.out
start time=2021-06-01T01:33:27Z flow cell id=FAO40255 protocol group id=210531 FAO40255 Pan 21s002174
sample id=Pan 21s002174
A TAAGATGTAAGACATGAATGCACATAAGATTATTTATAGTAACTAAACATATTGGATTAGTCACTGCACGGAATGATGTTAGCAATGTCTTCCTGGTTTTGAAATCTTAAA
{f TGAACTACAAGAGAAGGTGATCCAGCTACAGAGAGGTGATATTGATAATTGATCATAAATTGATCAACAAATTATCACCGATTTTTAATTTTAAGATTTCAAAAA
TAAAATACGCTGTTATATTCCATTTGCTGCGTGTGTTTAACCAATATATTTAATTTATATTTACTAAAATAATCTTATATTACATTCCGCATTTCTTTTGCTACATTAGCTA
%%),.1/--))+*+'&%($%(#$(&%%&54969<>>>;>983)7==>AHAOI:@DA6...,,,*****)))(((((('%/+7<B6,''&&&&&&'%)-23-
336696..*59179:=>D@CFG=CAD@>;;9&$)00+),-
8;;557<9:84<>1+6@?=9>E@?<11()&,.;A8=<75)1%&&20/02+(.%.09:;7D7DJDEHF<==5'%&)5=3,244:6;>0151,/0027;:9><(''*/&%&&&)
) () *) *) %##%/00(7.) *.5667-
32))688:CAD@:(78,7,4754(''),((1***./)$$'4016.,64=95<;;79>*+**.0+82:76.''78:>==4312/$'(28*)/97;88=:<;<)(*-,'5,44(
(,/.+6++,)(./+1::@B@9?5ACEDF@<C54.5=9<:8>?CB4CGGKHB9-''(4:@HE;6.,--
(-,1113:0??A?C>F;BG950??009/%#'.1322/)*01.8;>>?=67>:0FCB>B:;55*6;?AEE??B804?*'2.54%494503-0111.(4)%&&&*2463*-
48<9+',%)($%+%(:.$$1$%379:53*++'$,-.''(',2:8)8?@98<>5,6;60/-&%%.3774++,-/*))+'(*-
/13539@=673598;<C;A?=;42341.7264469:>HDA>C@7;;?;7;;8:561)0:;6<=>88:6%*/-)&&8+41??=3657:/1('$+%'%&%,691%+*(/78:?
@@AB@A9.618+$%')/681-,*3501134++*+,%+*'$(''')($$'#$#$&''('('%&&'$$'%%%%$$$$$$$$$##%#$$$&%)()$$'$%+)%.'*'%%%$)+,
```

++%) α *) α % α \$% α **) α \$% α **) α \$%) α **) α %) α **) α **) α ** α **