**SUPPLEMENTARY DATA**

**Supplemental Data 1 | K-Ras analysis forward primer**

1. library**(**"sangerseqR"**)**
2. library**(**"stringr"**)**
3. #setwd("N:/Documenten/Labjournal/2018-06-21 K-Ras seq")
4. path **=** "N:/documenten/sequences/180628NE-071/FW/"
5. out.file**<-**""
6. file.names **<-** dir**(**path, pattern **=**".ab1"**)**
7. **for(**k **in** 1**:**length**(**file.names**)){**
8. hetsangerseq **<-** readsangerseq**(**file.names**[**k**])**
9. trace **<-** traceMatrix**(**hetsangerseq**)**
10. peakpos **<-** peakPosMatrix**(**hetsangerseq**)**
11. # GTGGTTGGAGCT are bases prior to mutation
12. match\_pos**=**str\_locate**(**primarySeq**(**hetsangerseq, string **= TRUE)**, "TACGCCAC"**)** #reverse match=GTGGTTGGAGCT
13. #takes last position (e.g. 167) add 1 (168)
14. **if(!**is.na**(**match\_pos**[**2**])){**

15

16 mutation\_pos**=**match\_pos**[**2**]+**1 17

18 print **(**mutation\_pos**)**

19

1. #locate position of all bases
2. match\_T **<-** str\_locate\_all**(**primarySeq**(**hetsangerseq, string **= TRUE)**, "T"**)** 22 match\_T **<-** match\_T**[[**1**]][**,1**]**

23 match\_C **<-** str\_locate\_all**(**primarySeq**(**hetsangerseq, string **= TRUE)**, "C"**)** 24 match\_C **<-** match\_C**[[**1**]][**,1**]**

25

26

1. #we are looking for C to T mutations thus find nearest C and T peaks in sequence
2. #we add one to the match\_pos since this is where the mutation lies, and we want the next match in T base

29

1. sum\_T\_col **<-**""
2. **for(**i **in** 1**:**length**(**match\_T**)){**
3. **if(**match\_T**[**i**]>**mutation\_pos**){**
4. **if (**match\_T**[**i**-**1**]==**mutation\_pos**){**
5. selected\_T\_pos **<-** c**(**match\_T**[**i**-**5**]**,match\_T**[**i**-**4**]**, match\_T**[**i**-**3**]**, match\_T**[**i**-**2**]**, mutation\_pos, match\_T**[**i**]**, match\_T**[**i**+**1**]**, match\_T**[**i**+**2**]**, match\_T**[**i**+**3**])**
6. **} else {**
7. selected\_T\_pos **<-** c**(**match\_T**[**i**-**4**]**,match\_T**[**i**-**3**]**, match\_T**[**i**-**2**]**, match\_T**[**i**-**1**]**, mutation\_pos, match\_T**[**i**]**, match\_T**[**i**+**1**]**, match\_T**[**i**+**2**]**, match\_T**[**i**+**3**])**

37 **}**

38 **break**

39 **}**

40 **}**

41 **for (**j **in** selected\_T\_pos**){** 42

1. #peak trace column 1 = A
2. sum\_T **<-** sum**(**trace**[(**peakpos**[**j**]-**4**):(**peakpos**[**j**]+**4**)**,4**])** #1=A, 4=T
3. sum\_T\_col **<-** rbind**(**sum\_T\_col, sum\_T**)** 46 **}**

47

1. #we are looking for C to T mutations thus find nearest C and T peaks in sequence
2. #we add one to the match\_pos since this is where the mutation lies, and we want the next match in C base

50

51

1. **for(**i **in** 1**:**length**(**match\_C**)){**
2. **if(**match\_C**[**i**]>**mutation\_pos**){**
3. **if (**match\_C**[**i**-**1**]==**mutation\_pos**){**
4. selected\_C\_pos **<-** c**(**match\_C**[**i**-**5**]**,match\_C**[**i**-**4**]**, match\_C**[**i**-**3**]**, match\_C**[**i**-**2**]**, mutation\_pos, match\_C**[**i**]**, match\_C**[**i**+**1**]**, match\_C**[**i**+**2**]**, match\_C**[**i**+**3**])**
5. **} else {**
6. selected\_C\_pos **<-** c**(**match\_C**[**i**-**4**]**,match\_C**[**i**-**3**]**, match\_C**[**i**-**2**]**, match\_C**[**i**-**1**]**, mutation\_pos, match\_C**[**i**]**, match\_C**[**i**+**1**]**, match\_C**[**i**+**2**]**, match\_C**[**i**+**3**])**

58 **}**

59

60 **break**

61 **}**

62 **}**

63

1. sum\_C\_col **<-** ""
2. **for (**j **in** selected\_C\_pos**){**
3. #peak trace column 1 = G
4. sum\_C **<-**sum**(**trace**[(**peakpos**[**j**]-**4**):(**peakpos**[**j**]+**4**)**,2**])** #3=G, 2=C
5. sum\_C\_col **<-** rbind**(**sum\_C\_col, sum\_C**)** 70

71 **}**

1. sum\_bases **<-** c**(**sum\_T\_col, sum\_C\_col**)**
2. write**(**file.names**[**k**]**, file **=** "K-RAS\_peaks.txt", append **= TRUE**, sep**=**","**)**
3. write**(**sum\_bases, file **=** "K-RAS\_peaks.txt", append **= TRUE**, sep**=**","**)** 75 **}**

76 **}**

77

**Supplemental Data 2 | K-Ras analysis reverse primer**

1. library**(**"sangerseqR"**)**
2. library**(**"stringr"**)**
3. path **=** "N:/documenten/sequences/180628NE-071/RV/"
4. out.file**<-**""
5. file.names **<-** dir**(**path, pattern **=**".ab1"**)**
6. **for(**k **in** 1**:**length**(**file.names**)){**
7. hetsangerseq **<-** readsangerseq**(**file.names**[**k**])**
8. trace **<-** traceMatrix**(**hetsangerseq**)**
9. peakpos **<-** peakPosMatrix**(**hetsangerseq**)**
10. # GTGGTTGGAGCT are bases prior to mutation
11. match\_pos**=**str\_locate**(**primarySeq**(**hetsangerseq, string **= TRUE)**, "GTGGTTGGAGCT"**)**
12. #takes last position (e.g. 167) add 1 (168)
13. **if(!**is.na**(**match\_pos**[**2**])){**
14. mutation\_pos**=**match\_pos**[**2**]+**1 15
15. #locate position of all bases
16. match\_A **<-** str\_locate\_all**(**primarySeq**(**hetsangerseq, string **= TRUE)**, "A"**)** 18 match\_A **<-** match\_A**[[**1**]][**,1**]**

19 match\_G **<-** str\_locate\_all**(**primarySeq**(**hetsangerseq, string **= TRUE)**, "G"**)** 20 match\_G **<-** match\_G**[[**1**]][**,1**]**

21

22

1. #we are looking for G to A mutations thus find nearest G and A peaks in sequence
2. #we add one to the match\_pos since this is where the mutation lies, and we want the next match in A base

25

1. sum\_A\_col **<-**""
2. **for(**i **in** 1**:**length**(**match\_A**)){**
3. **if(**match\_A**[**i**]>**mutation\_pos**){**
4. **if (**match\_A**[**i**-**1**]==**mutation\_pos**){**
5. selected\_A\_pos **<-** c**(**match\_A**[**i**-**5**]**,match\_A**[**i**-**4**]**, match\_A**[**i**-**3**]**, match\_A**[**i**-**2**]**, mutation\_pos, match\_A**[**i**]**, match\_A**[**i**+**1**]**, match\_A**[**i**+**2**]**, match\_A**[**i**+**3**])**
6. **} else {**
7. selected\_A\_pos **<-** c**(**match\_A**[**i**-**4**]**,match\_A**[**i**-**3**]**, match\_A**[**i**-**2**]**, match\_A**[**i**-**1**]**, mutation\_pos, match\_A**[**i**]**, match\_A**[**i**+**1**]**, match\_A**[**i**+**2**]**, match\_A**[**i**+**3**])**

33 **}**

34 **break**

35 **}**

36 **}**

37 **for (**j **in** selected\_A\_pos**){** 38

1. #peak trace column 1 = A
2. sum\_A **<-** sum**(**trace**[(**peakpos**[**j**]-**4**):(**peakpos**[**j**]+**4**)**,1**])**
3. sum\_A\_col **<-** rbind**(**sum\_A\_col, sum\_A**)** 42 **}**

43

1. #we are looking for G to A mutations thus find nearest G and A peaks in sequence
2. #we add one to the match\_pos since this is where the mutation lies, and we want the next match in G base

46

47

1. **for(**i **in** 1**:**length**(**match\_G**)){**
2. **if(**match\_G**[**i**]>**mutation\_pos**){**
3. **if (**match\_G**[**i**-**1**]==**mutation\_pos**){**
4. selected\_G\_pos **<-** c**(**match\_G**[**i**-**5**]**,match\_G**[**i**-**4**]**, match\_G**[**i**-**3**]**, match\_G**[**i**-**2**]**, mutation\_pos, match\_G**[**i**]**, match\_G**[**i**+**1**]**, match\_G**[**i**+**2**]**, match\_G**[**i**+**3**])**
5. **} else {**
6. selected\_G\_pos **<-** c**(**match\_G**[**i**-**4**]**,match\_G**[**i**-**3**]**, match\_G**[**i**-**2**]**, match\_G**[**i**-**1**]**, mutation\_pos, match\_G**[**i**]**, match\_G**[**i**+**1**]**, match\_G**[**i**+**2**]**, match\_G**[**i**+**3**])**

54 **}**

55

56 **break**

57 **}**

58 **}**

59

1. sum\_G\_col **<-** ""
2. **for (**j **in** selected\_G\_pos**){** 62
3. #peak trace column 1 = G
4. sum\_G **<-**sum**(**trace**[(**peakpos**[**j**]-**4**):(**peakpos**[**j**]+**4**)**,3**])**
5. sum\_G\_col **<-** rbind**(**sum\_G\_col, sum\_G**)** 66

67 **}**

1. sum\_bases **<-** c**(**sum\_A\_col, sum\_G\_col**)**
2. write**(**file.names**[**k**]**, file **=** "K-RAS\_peaks.txt", append **= TRUE**, sep**=**","**)**
3. write**(**sum\_bases, file **=** "K-RAS\_peaks.txt", append **= TRUE**, sep**=**","**)** 71 **}**

72 **}**

73

74