**21 August 2014**

**In silico modeling of performance of DNA mixes**

Refer to excel file “140819\_Final SequIn Mixes\_DNAv2.xlsx”.

**Mix A**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **REF** | **VAR** | **Grp** | **REF length** | **VAR length** | **AIM** | **RATIO** | **RATIO** | **CON** | **CON** | **CON READ** | **CON READS** | **PER KB** | **PER KB** | **ROUND** | **ROUND** |
| D\_1\_6\_R | D\_1\_6\_V | C | 1024 | 1000 | 2.00E+07 | 1 | 1 | 0.5 | 0.5 | 10000000 | 10000000 | 10240000 | 10000000 | 10240000 | 10000000 |
| D\_2\_6\_R | D\_2\_6\_V | C | 978 | 969 | 2.00E+07 | 1 | 9 | 0.1 | 0.9 | 2000000 | 18000000 | 1956000 | 17442000 | 1956000 | 17442000 |
| D\_3\_6\_R | D\_3\_6\_V | C | 1079 | 1104 | 2.00E+07 | 1 | 99 | 0.01 | 0.99 | 200000 | 19800000 | 215800 | 21859200 | 215800 | 21859200 |
| D\_1\_10\_R | D\_1\_10\_V | B | 927 | 900 | 6.67E+06 | 1 | 1 | 0.5 | 0.5 | 3333333 | 3333333 | 3089999.691 | 2999999.7 | 3089999 | 2999999 |
| D\_2\_10\_R | D\_2\_10\_V | B | 1004 | 1029 | 6.67E+06 | 9 | 1 | 0.9 | 0.1 | 5999999.4 | 666666.6 | 6023999.398 | 685999.9314 | 6023999 | 685999 |
| D\_3\_9\_R | D\_3\_9\_V | B | 1007 | 1027 | 6.67E+06 | 99 | 1 | 0.99 | 0.01 | 6599999.34 | 66666.66 | 6646199.335 | 68466.65982 | 6646199 | 68466 |
| D\_1\_11\_R | D\_1\_11\_V | D | 1115 | 1121 | 2222222 | 1 | 1 | 0.5 | 0.5 | 1111111 | 1111111 | 1238888.765 | 1245555.431 | 1238888 | 1245555 |
| D\_2\_11\_R | D\_2\_11\_V | D | 1039 | 1027 | 2222222 | 1 | 9 | 0.1 | 0.9 | 222222.2 | 1999999.8 | 230888.8658 | 2053999.795 | 230888 | 2053999 |
| D\_3\_11\_R | D\_3\_11\_V | D | 953 | 969 | 2222222 | 1 | 99 | 0.01 | 0.99 | 22222.22 | 2199999.78 | 21177.77566 | 2131799.787 | 21177 | 2131799 |
| D\_1\_12\_R | D\_1\_12\_V | A | 979 | 997 | 740740 | 1 | 1 | 0.5 | 0.5 | 370370 | 370370 | 362592.23 | 369258.89 | 362592 | 369258 |
| D\_2\_12\_R | D\_2\_12\_V | A | 994 | 973 | 740740 | 9 | 1 | 0.9 | 0.1 | 666666 | 74074 | 662666.004 | 72074.002 | 662666 | 72074 |
| D\_3\_12\_R | D\_3\_12\_V | A | 1002 | 980 | 740740 | 99 | 1 | 0.99 | 0.01 | 733332.6 | 7407.4 | 734799.2652 | 7259.252 | 734799 | 7259 |
| D\_1\_2\_R | D\_1\_2\_V | D | 1122 | 1134 | 246912 | 1 | 1 | 0.5 | 0.5 | 123456 | 123456 | 138517.632 | 139999.104 | 138517 | 139999 |
| D\_2\_2\_R | D\_2\_2\_V | D | 959 | 978 | 246913 | 1 | 9 | 0.1 | 0.9 | 24691.3 | 222221.7 | 23678.9567 | 217332.8226 | 23678 | 217332 |
| D\_2\_9\_R | D\_2\_9\_V | D | 973 | 989 | 246913 | 1 | 99 | 0.01 | 0.99 | 2469.13 | 244443.87 | 2402.46349 | 241754.9874 | 2402 | 241754 |
| D\_1\_3\_R | D\_1\_3\_V | D | 986 | 983 | 82304 | 1 | 1 | 0.5 | 0.5 | 41152 | 41152 | 40575.872 | 40452.416 | 40575 | 40452 |
| D\_1\_9\_R | D\_1\_9\_V | D | 1099 | 1077 | 82304 | 9 | 1 | 0.9 | 0.1 | 74073.6 | 8230.4 | 81406.8864 | 8864.1408 | 81406 | 8864 |
| D\_3\_3\_R | D\_3\_3\_V | D | 1018 | 1034 | 82304 | 99 | 1 | 0.99 | 0.01 | 81480.96 | 823.04 | 82947.61728 | 851.02336 | 82947 | 851 |
| D\_1\_4\_R | D\_1\_4\_V | C | 1115 | 1103 | 27434 | 1 | 1 | 0.5 | 0.5 | 13717 | 13717 | 15294.455 | 15129.851 | 15294 | 15129 |
| D\_2\_4\_R | D\_2\_4\_V | C | 968 | 968 | 27434 | 1 | 9 | 0.1 | 0.9 | 2743.4 | 24690.6 | 2655.6112 | 23900.5008 | 2655 | 23900 |
| D\_3\_4\_R | D\_3\_4\_V | C | 1096 | 1099 | 27434 | 1 | 99 | 0.01 | 0.99 | 274.34 | 27159.66 | 300.67664 | 29848.46634 | 300 | 29848 |
| D\_1\_5\_R | D\_1\_5\_V | B | 943 | 939 | 9144 | 1 | 1 | 0.5 | 0.5 | 4572 | 4572 | 4311.396 | 4293.108 | 4311 | 4293 |
| D\_2\_5\_R | D\_2\_5\_V | B | 1121 | 1119 | 9144 | 9 | 1 | 0.9 | 0.1 | 8229.6 | 914.4 | 9225.3816 | 1023.2136 | 9225 | 1023 |
| D\_3\_5\_R | D\_3\_5\_V | B | 1130 | 1135 | 9144 | 99 | 1 | 0.99 | 0.01 | 9052.56 | 91.44 | 10229.3928 | 103.7844 | 10229 | 103 |
| D\_1\_1\_R | D\_1\_1\_V | D | 1107 | 1133 | 3048 | 1 | 1 | 0.5 | 0.5 | 1524 | 1524 | 1687.068 | 1726.692 | 1687 | 1726 |
| D\_2\_1\_R | D\_2\_1\_V | D | 935 | 958 | 3047 | 1 | 9 | 0.1 | 0.9 | 304.7 | 2742.3 | 284.8945 | 2627.1234 | 284 | 2627 |
| D\_3\_1\_R | D\_3\_1\_V | D | 934 | 945 | 3047 | 1 | 99 | 0.01 | 0.99 | 30.47 | 3016.53 | 28.45898 | 2850.62085 | 28 | 2850 |
| D\_1\_7\_R | D\_1\_7\_V | A | 1057 | 1062 | 1016 | 1 | 1 | 0.5 | 0.5 | 508 | 508 | 536.956 | 539.496 | 536 | 539 |
| D\_2\_7\_R | D\_2\_7\_V | A | 946 | 950 | 1015 | 9 | 1 | 0.9 | 0.1 | 913.5 | 101.5 | 864.171 | 96.425 | 864 | 96 |
| D\_3\_7\_R | D\_3\_7\_V | A | 893 | 905 | 1015 | 99 | 1 | 0.99 | 0.01 | 1004.85 | 10.15 | 897.33105 | 9.18575 | 897 | 9 |
| D\_1\_8\_R | D\_1\_8\_V | D | 1118 | 1125 | 338 | 1 | 1 | 0.5 | 0.5 | 169 | 169 | 188.942 | 190.125 | 188 | 190 |
| D\_2\_8\_R | D\_2\_8\_V | D | 1044 | 1049 | 337 | 1 | 9 | 0.1 | 0.9 | 33.7 | 303.3 | 35.1828 | 318.1617 | 35 | 318 |
| D\_3\_8\_R | D\_3\_8\_V | D | 1065 | 1087 | 338 | 1 | 99 | 0.01 | 0.99 | 3.38 | 334.62 | 3.5997 | 363.73194 | 3 | 363 |
| D\_2\_3\_R | D\_2\_3\_V | D | 936 | 938 | 112 | 1 | 1 | 0.5 | 0.5 | 56 | 56 | 52.416 | 52.528 | 52 | 52 |
| D\_3\_2\_R | D\_3\_2\_V | D | 793 | 783 | 112 | 9 | 1 | 0.9 | 0.1 | 100.8 | 11.2 | 79.9344 | 8.7696 | 79 | 8 |
| D\_3\_10\_R | D\_3\_10\_V | D | 925 | 941 | 112 | 99 | 1 | 0.99 | 0.01 | 110.88 | 1.12 | 102.564 | 1.05392 | 102 | 1 |

**Mix B**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **REF** | **VAR** | **Grp** | **REF length** | **VAR length** | **AIM** | **RATIO** | **RATIO** | **CON** | **CON** | **CON READ** | **CON READS** | **PER KB** | **PER KB** | **ROUND** | **ROUND** |
| D\_1\_6\_R | D\_1\_6\_V | C | 1024 | 1000 | 20000000 | 1 | 1 | 0.5 | 0.5 | 10000000 | 10000000 | 10240000 | 10000000 | 10240000 | 10000000 |
| D\_2\_6\_R | D\_2\_6\_V | C | 978 | 969 | 20000000 | 9 | 1 | 0.9 | 0.1 | 18000000 | 2000000 | 17604000 | 1938000 | 17604000 | 1938000 |
| D\_3\_6\_R | D\_3\_6\_V | C | 1079 | 1104 | 20000000 | 99 | 1 | 0.99 | 0.01 | 19800000 | 200000 | 21364200 | 220800 | 21364200 | 220800 |
| D\_1\_10\_R | D\_1\_10\_V | B | 927 | 900 | 2222222 | 1 | 1 | 0.5 | 0.5 | 1111111 | 1111111 | 1029999.897 | 999999.9 | 1029999 | 999999 |
| D\_2\_10\_R | D\_2\_10\_V | B | 1004 | 1029 | 2222222 | 1 | 9 | 0.1 | 0.9 | 222222.2 | 1999999.8 | 223111.0888 | 2057999.794 | 223111 | 2057999 |
| D\_3\_9\_R | D\_3\_9\_V | B | 1007 | 1027 | 2222222 | 1 | 99 | 0.01 | 0.99 | 22222.22 | 2199999.78 | 22377.77554 | 2259399.774 | 22377 | 2259399 |
| D\_1\_11\_R | D\_1\_11\_V | D | 1115 | 1121 | 6666666 | 1 | 1 | 0.5 | 0.5 | 3333333 | 3333333 | 3716666.295 | 3736666.293 | 3716666 | 3736666 |
| D\_2\_11\_R | D\_2\_11\_V | D | 1039 | 1027 | 6666666 | 9 | 1 | 0.9 | 0.1 | 5999999.4 | 666666.6 | 6233999.377 | 684666.5982 | 6233999 | 684666 |
| D\_3\_11\_R | D\_3\_11\_V | D | 953 | 969 | 6666666 | 99 | 1 | 0.99 | 0.01 | 6599999.34 | 66666.66 | 6289799.371 | 64599.99354 | 6289799 | 64599 |
| D\_1\_12\_R | D\_1\_12\_V | A | 979 | 997 | 82304 | 1 | 1 | 0.5 | 0.5 | 41152 | 41152 | 40287.808 | 41028.544 | 40287 | 41028 |
| D\_2\_12\_R | D\_2\_12\_V | A | 994 | 973 | 82304 | 1 | 9 | 0.1 | 0.9 | 8230.4 | 74073.6 | 8181.0176 | 72073.6128 | 8181 | 72073 |
| D\_3\_12\_R | D\_3\_12\_V | A | 1002 | 980 | 82304 | 1 | 99 | 0.01 | 0.99 | 823.04 | 81480.96 | 824.68608 | 79851.3408 | 824 | 79851 |
| D\_1\_2\_R | D\_1\_2\_V | D | 1122 | 1134 | 740740 | 1 | 1 | 0.5 | 0.5 | 370370 | 370370 | 415555.14 | 419999.58 | 415555 | 419999 |
| D\_2\_2\_R | D\_2\_2\_V | D | 959 | 978 | 740740 | 9 | 1 | 0.9 | 0.1 | 666666 | 74074 | 639332.694 | 72444.372 | 639332 | 72444 |
| D\_2\_9\_R | D\_2\_9\_V | D | 973 | 989 | 740740 | 99 | 1 | 0.99 | 0.01 | 733332.6 | 7407.4 | 713532.6198 | 7325.9186 | 713532 | 7325 |
| D\_1\_3\_R | D\_1\_3\_V | D | 986 | 983 | 246912 | 1 | 1 | 0.5 | 0.5 | 123456 | 123456 | 121727.616 | 121357.248 | 121727 | 121357 |
| D\_1\_9\_R | D\_1\_9\_V | D | 1099 | 1077 | 246913 | 1 | 9 | 0.1 | 0.9 | 24691.3 | 222221.7 | 27135.7387 | 239332.7709 | 27135 | 239332 |
| D\_3\_3\_R | D\_3\_3\_V | D | 1018 | 1034 | 246913 | 1 | 99 | 0.01 | 0.99 | 2469.13 | 244443.87 | 2513.57434 | 252754.9616 | 2513 | 252754 |
| D\_1\_4\_R | D\_1\_4\_V | C | 1115 | 1103 | 27434 | 1 | 1 | 0.5 | 0.5 | 13717 | 13717 | 15294.455 | 15129.851 | 15294 | 15129 |
| D\_2\_4\_R | D\_2\_4\_V | C | 968 | 968 | 27434 | 9 | 1 | 0.9 | 0.1 | 24690.6 | 2743.4 | 23900.5008 | 2655.6112 | 23900 | 2655 |
| D\_3\_4\_R | D\_3\_4\_V | C | 1096 | 1099 | 27434 | 99 | 1 | 0.99 | 0.01 | 27159.66 | 274.34 | 29766.98736 | 301.49966 | 29766 | 301 |
| D\_1\_5\_R | D\_1\_5\_V | B | 943 | 939 | 3048 | 1 | 1 | 0.5 | 0.5 | 1524 | 1524 | 1437.132 | 1431.036 | 1437 | 1431 |
| D\_2\_5\_R | D\_2\_5\_V | B | 1121 | 1119 | 3047 | 1 | 9 | 0.1 | 0.9 | 304.7 | 2742.3 | 341.5687 | 3068.6337 | 341 | 3068 |
| D\_3\_5\_R | D\_3\_5\_V | B | 1130 | 1135 | 3047 | 1 | 99 | 0.01 | 0.99 | 30.47 | 3016.53 | 34.4311 | 3423.76155 | 34 | 3423 |
| D\_1\_1\_R | D\_1\_1\_V | D | 1107 | 1133 | 9144 | 1 | 1 | 0.5 | 0.5 | 4572 | 4572 | 5061.204 | 5180.076 | 5061 | 5180 |
| D\_2\_1\_R | D\_2\_1\_V | D | 935 | 958 | 9144 | 9 | 1 | 0.9 | 0.1 | 8229.6 | 914.4 | 7694.676 | 875.9952 | 7694 | 875 |
| D\_3\_1\_R | D\_3\_1\_V | D | 934 | 945 | 9144 | 99 | 1 | 0.99 | 0.01 | 9052.56 | 91.44 | 8455.09104 | 86.4108 | 8455 | 86 |
| D\_1\_7\_R | D\_1\_7\_V | A | 1057 | 1062 | 112 | 1 | 1 | 0.5 | 0.5 | 56 | 56 | 59.192 | 59.472 | 59 | 59 |
| D\_2\_7\_R | D\_2\_7\_V | A | 946 | 950 | 112 | 1 | 9 | 0.1 | 0.9 | 11.2 | 100.8 | 10.5952 | 95.76 | 10 | 95 |
| D\_3\_7\_R | D\_3\_7\_V | A | 893 | 905 | 112 | 1 | 99 | 0.01 | 0.99 | 1.12 | 110.88 | 1.00016 | 100.3464 | 1 | 100 |
| D\_1\_8\_R | D\_1\_8\_V | D | 1118 | 1125 | 1016 | 1 | 1 | 0.5 | 0.5 | 508 | 508 | 567.944 | 571.5 | 567 | 571 |
| D\_2\_8\_R | D\_2\_8\_V | D | 1044 | 1049 | 1015 | 9 | 1 | 0.9 | 0.1 | 913.5 | 101.5 | 953.694 | 106.4735 | 953 | 106 |
| D\_3\_8\_R | D\_3\_8\_V | D | 1065 | 1087 | 1015 | 99 | 1 | 0.99 | 0.01 | 1004.85 | 10.15 | 1070.16525 | 11.03305 | 1070 | 11 |
| D\_2\_3\_R | D\_2\_3\_V | D | 936 | 938 | 338 | 1 | 1 | 0.5 | 0.5 | 169 | 169 | 158.184 | 158.522 | 158 | 158 |
| D\_3\_2\_R | D\_3\_2\_V | D | 793 | 783 | 337 | 1 | 9 | 0.1 | 0.9 | 33.7 | 303.3 | 26.7241 | 237.4839 | 26 | 237 |
| D\_3\_10\_R | D\_3\_10\_V | D | 925 | 941 | 338 | 1 | 99 | 0.01 | 0.99 | 3.38 | 334.62 | 3.1265 | 314.87742 | 3 | 314 |

Create simulated reads with Sherman (samtools)

$ module load samtools/0.1.18

$ while read X

$ do

$ wgsim -d 400 -N 88000000 -1 101 -2 101 ${X} ${X}.R1.fq ${X}.R2.fq

$ done < /hox/u/uqtmerce/Wendy/DNA/mixes /DNA.standards.fa.list.txt

#where DNA.standards.fa.list.txt is a list of all the DNA standards in .fa, i.e.,

D\_10\_1\_R.fa

D\_10\_1\_V.fa

D\_10\_2\_R.fa

D\_10\_2\_V.fa

D\_10\_3l\_R.fa

D\_10\_3\_V.fa

…

#Create bowtie2 index files

bowtie2-build -f /hox/u/uqtmerce/Wendy/ChrT.5.10.fa ChrT.5.10

# Aliqn your simulated reads to the reference genome

$ while read X

$ do

$ bowtie2 -x ChrT.5.10 -1 ${X}.fa.R1.fq -2 ${X}.fa.R2.fq -S ${X}.sam

$ done < DNA.standards.list.txt

# where DNA.standards.list.txt is a list of all the DNA standards, i.e.,

D\_10\_1\_R

D\_10\_1\_V

D\_10\_2\_R

D\_10\_2\_V

D\_10\_3\_R

D\_10\_3\_V

…

$ head -3515000 sams/D\_1\_1\_R.sam > tempA/D\_1\_1\_R.mixA.sam

# do for all the .sam files

$ cat \*.sam > combined.sam

#alt mtd to get reads for variant

sed 's/\_/\t/4' combined.sam > temp.txt

makeHistogram.pl temp.txt 1

In the ideal situation, the mixes will appear as…

 

 

 