

# VCF ファイルの genotype の妥当性

VCF ファイルの genotype の信頼性に疑問があるので、以下の基準にしたがって判定した genotype に修正する。

## 1. BAM ファイルのリードから genotype を判定する基準

### (1) 判定基準( $Q1 \geq Q2$ が前提)

Q1 …quality がいちばん大きい allele の quality

Q2 …quality が 2 番目に大きい allele の quality

#### ● Diploid の場合

(領域①)  $Q1 < Q2 \times 5$  かつ  $Q2 \geq 40$  のとき → Hetero

ただし  $Q2 - Q3 \leq 5$  のときは、2 番目の allele を N とし警告を付加する。

(領域②)  $Q1 \geq Q2 \times 5$  かつ  $Q1 \geq 40$  のとき → Homo

(領域③)  $Q1 < Q2 \times 5$  かつ  $Q2 < 40$  かつ  $Q1 \geq 40$  のとき → Q1 の allele だけ採用する(他方は N)

(領域④)  $Q1 < 40$  のとき → N/N

#### ● Haploid の場合

(領域①②③)  $Q1 \geq 40$  のとき → Q1 の allele を採用する

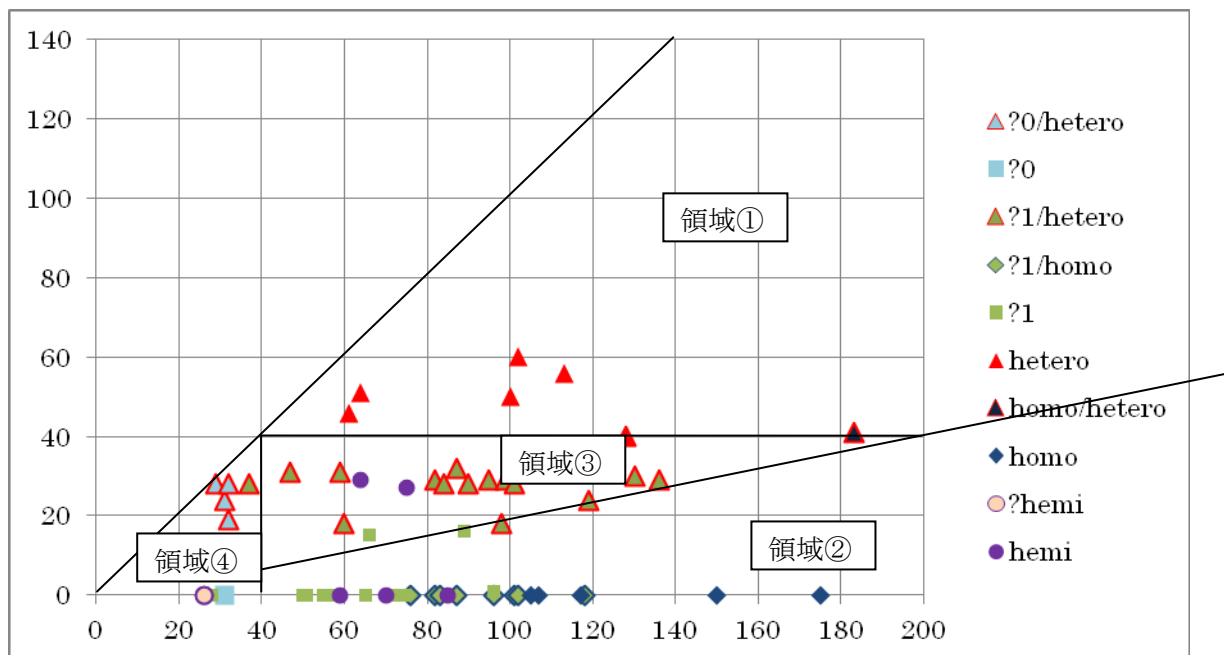
ただし  $Q1 - Q2 \leq 5$  のときは、N とし警告を付加する。

(領域④)  $Q1 < 40$  のとき → N

※ 領域の番号は(2)のグラフを参照

### (2) Quality と genotype 判定結果の関係(quality\_compare.xls 参照)

横軸=Q1, 縦軸=Q2



#### ● 凡例

?0 … まったく推定できない

?1 … 1 つは推定できるが 2 つは推定できない

?hemi … hemi であるが allele は推定できない

- “/”の左は「判定 1」、「/」の右は「判定 2 甘目」

(3) 考え方

- $Q < 40$  の allele は、存在するとは断定しない。(N にするか、他の allele の homo にする。)
- Quality の比が 5 以上のときは hetero ではない。(Homo か、N/N にする。)

## 2. 基準を検討する方法

VCF ファイルの genotype と BAM ファイルのリード情報を比較し、VCF ファイルの genotype が妥当であるかを判定した。

(1) 信頼性に疑問があるサイトの調査

Tree 上で孤立した場所に現れる個体から 5 個体を選び、変異サイトの Quality value を計算した。

VCF ファイルの genotype が homo であるサイトは、VCF ファイルが示す allele の Quality value と、それ以外の 3 つの allele の Quality value の最大値とを比較した。

VCF ファイルの genotype が hetero であるサイトは、VCF ファイルが示す 2 つの allele の Quality value を比較した。残りの 2 つの allele の Quality value も念のため確認した。

この結果、Quality value が低いのに VCF ファイルで call されていたり、Quality value が高いのに VCF ファイルで call されていないものがあるなど、整合しないサイトがあった。

当初はリードの base quality minimum  $\geq 26$  のフィルターを設定し、quality  $< 26$  のサイトは除いて (quality=0 として) Quality value を計算した (黒字)。その後に base quality minimum  $\geq 0$  に変更し (=最小値を定めずに) 改めて計算した (茶色字)。

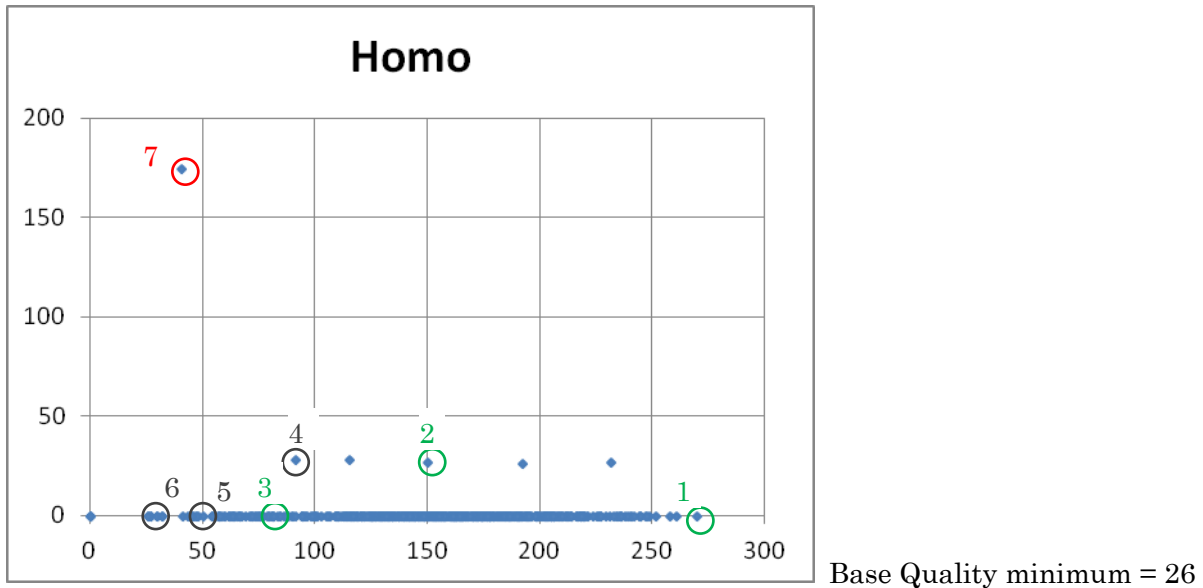
(2) Quality value と妥当な genotype の関係の調査

次に UCSC genome browser の画面を目視して判断した genotype と、VCF ファイルの genotype が一致しないサイトについて、UCSC genome browser の画面を目視して判断した genotype と BAM ファイルから計算した Quality value との関係調べた。これをもとにして、基準を設定した。

### 3. データ

① 個体名: NA19469

(a) VCF が Homo であるサイト



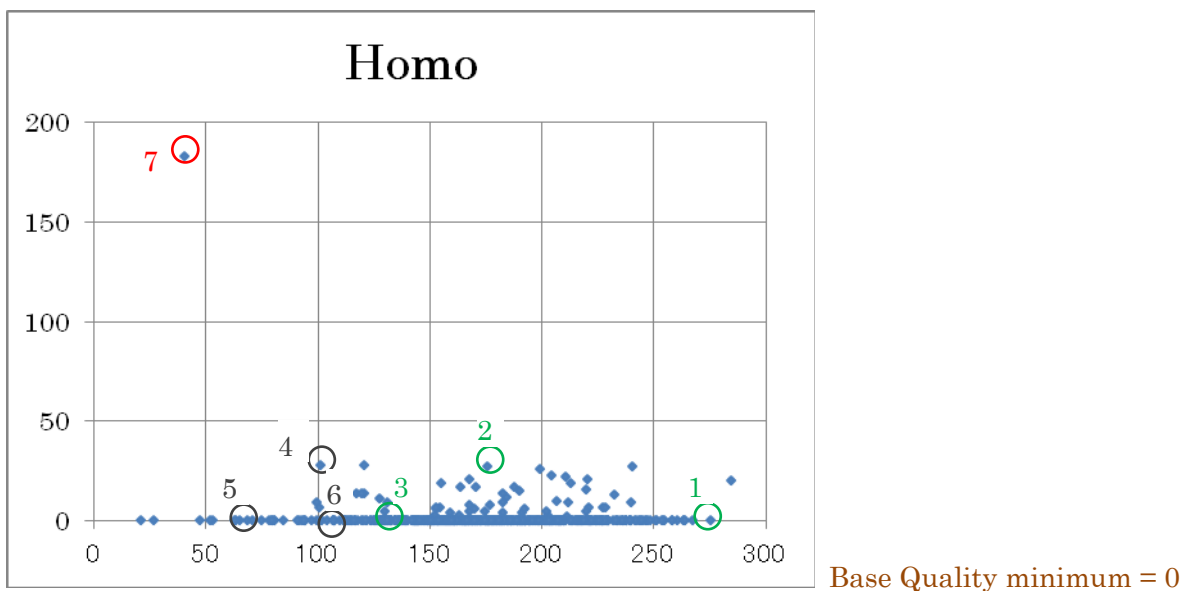
横軸: VCF が示す allele の Quality value

縦軸: それ以外の 3 つの allele の Quality value の最大値

1～3 は VCF ファイルの genotype が正しそうです。


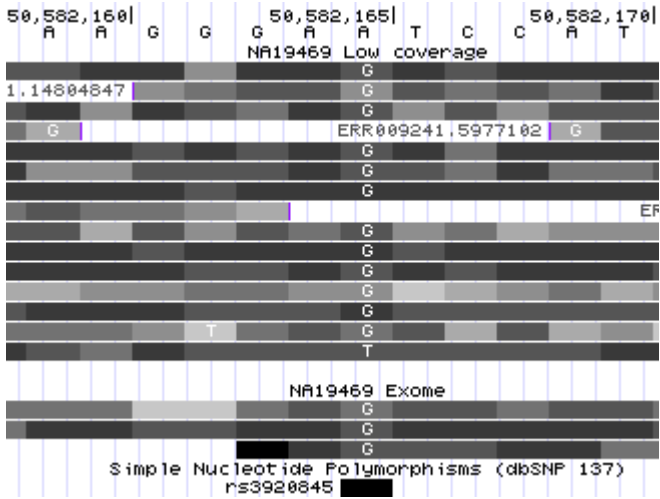
4～6 は VCF ファイルの genotype に矛盾は少ないが、リードの質が低いので判断が難しい。

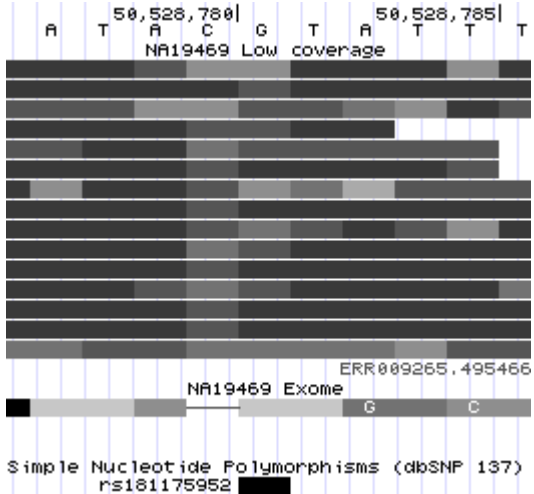
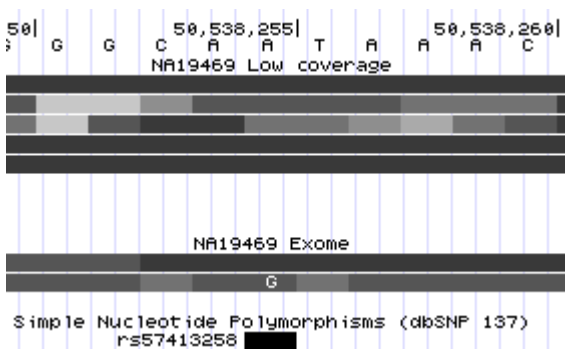
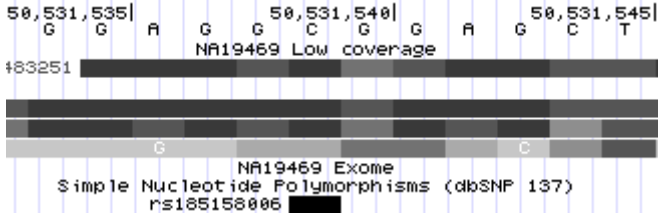
7 は VCF ファイルの genotype は正しくなさそうです。

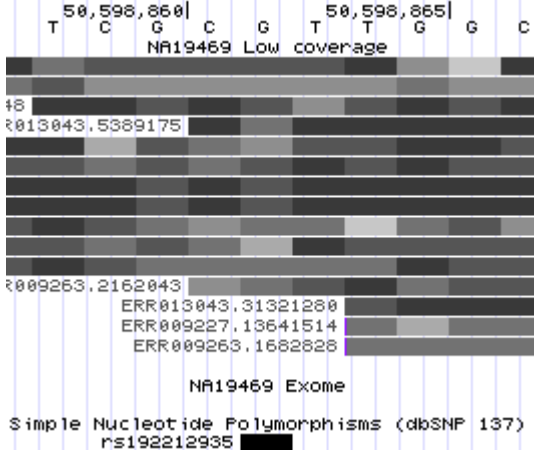
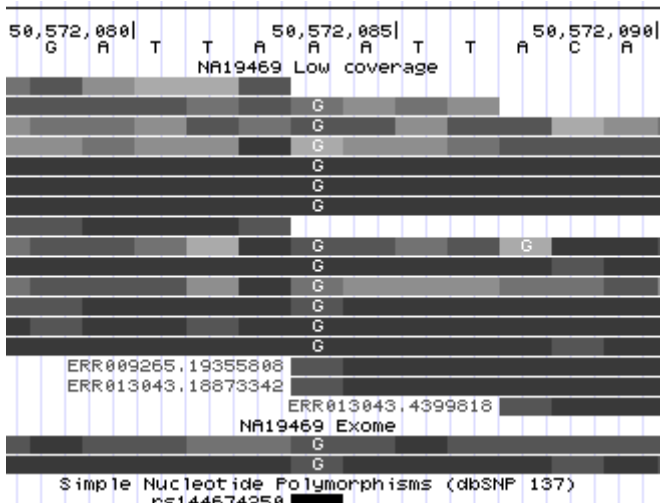


横軸: VCF が示す allele の Quality value

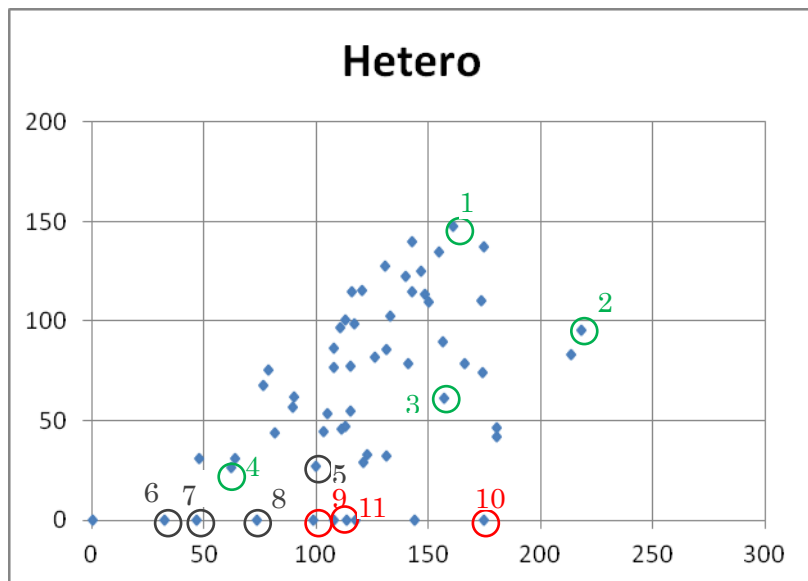
縦軸: それ以外の 3 つの allele の Quality value の最大値

	UCSC Genome browser	
1		<p>chrX:50569885, REF=C, ALT=A VCF genotype = C   C <math>Q_v(C)=270</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=270, 比=×</p> <p>正しい genotype: C   C (=VCF)</p> <p>Base Quality minimum = 0 のとき <math>Q_v(C)=275</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=275, 比=×</p>
2		<p>chrX:50582165, REF=A, ALT=G VCF genotype = G   G <math>Q_v(G)=150</math>, <math>Q_v(T)=27</math>, <math>Q_v(A)=Q_v(C)=0</math> 差=123, 比=5.6</p> <p>正しい genotype: G   G (=VCF)</p> <p>Base Quality minimum = 0 のとき <math>Q_v(G)=176</math>, <math>Q_v(T)=27</math>, <math>Q_v(A)=Q_v(C)=0</math> 差=149, 比=6.5</p>

3	 <p>50,528,780  50,528,785  A T A C G T A T T T NA19469 Low coverage ERR009265.495466 NA19469 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs181175952</p>	<p>chrX:50528781, REF=G, ALT=A VCF genotype = G   G <math>Q_v(G)=80</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=80, 比=×</p> <p>正しい genotype: G   G (VCF)</p> <p>Base Quality minimum = 0 のとき <math>Q_v(G)=133</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=133, 比=×</p>
4	 <p>50  50,538,255  50,538,260  G G C A A T A A A C NA19469 Low coverage NA19469 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs57413258</p>	<p>chrX:50538255, REF=A, ALT=G VCF genotype = A   A <math>Q_v(A)=91</math>, <math>Q_v(G)=28</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=63, 比=3.3</p> <p>正しい genotype: A/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(A)=101</math>, <math>Q_v(G)=28</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=73, 比=3.6</p>
5	 <p>50,531,535  50,531,540  50,531,545  G G A G G C G G A G C T NA19469 Low coverage 183251 NA19469 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs185158006</p>	<p>chrX:50531539, REF=C, ALT=T VCF genotype = C   C <math>Q_v(C)=55</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=55, 比=×</p> <p>正しい genotype: C/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(C)=65</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=65, 比=×</p>

6	 <p>chrX:50598862, REF=G, ALT=A VCF genotype = G   G <math>Q_v(G)=26</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=26, 比=×</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(G)=107</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=107, 比=×</p>	<p>chrX:50598862, REF=G, ALT=A VCF genotype = G   G <math>Q_v(G)=26</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=26, 比=×</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(G)=107</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=107, 比=×</p>
7	 <p>chrX:50572084, REF=A, ALT=G VCF genotype = A   A <math>Q_v(A)=41</math>, <math>Q_v(G)=175</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=-134, 比=0.23</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(A)=41</math>, <math>Q_v(G)=183</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=-142, 比=0.22</p>	<p>chrX:50572084, REF=A, ALT=G VCF genotype = A   A <math>Q_v(A)=41</math>, <math>Q_v(G)=175</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=-134, 比=0.23</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(A)=41</math>, <math>Q_v(G)=183</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=-142, 比=0.22</p>

(b) VCF が Hetero であるサイト



Base Quality minimum = 26

横軸: VCF が示す 2 つの allele の Quality value のうち大きい方

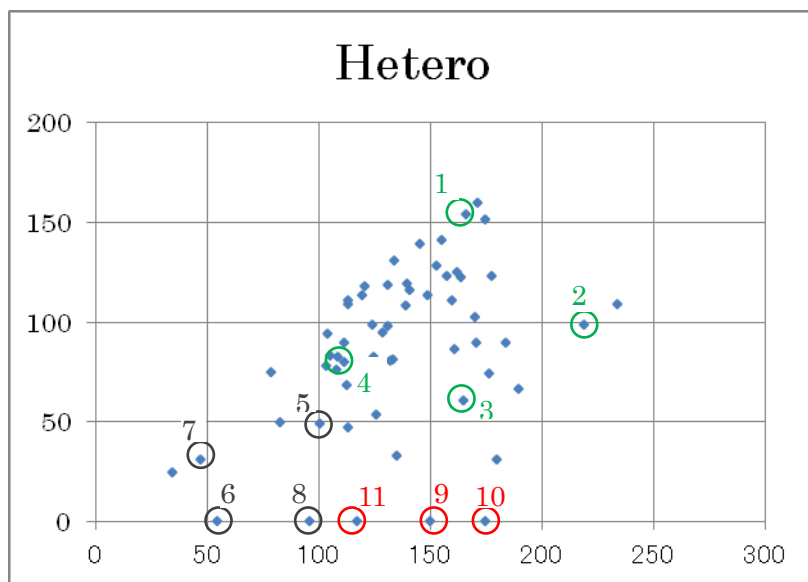
縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方

(VCF が示す allele 以外の残り 2 つはいずれも 0 であった。)

1~4 は VCF ファイルの genotype が正しそうである。

5~8 は VCF ファイルの genotype に矛盾は少ないが、リードの質が低いので判断が難しい。

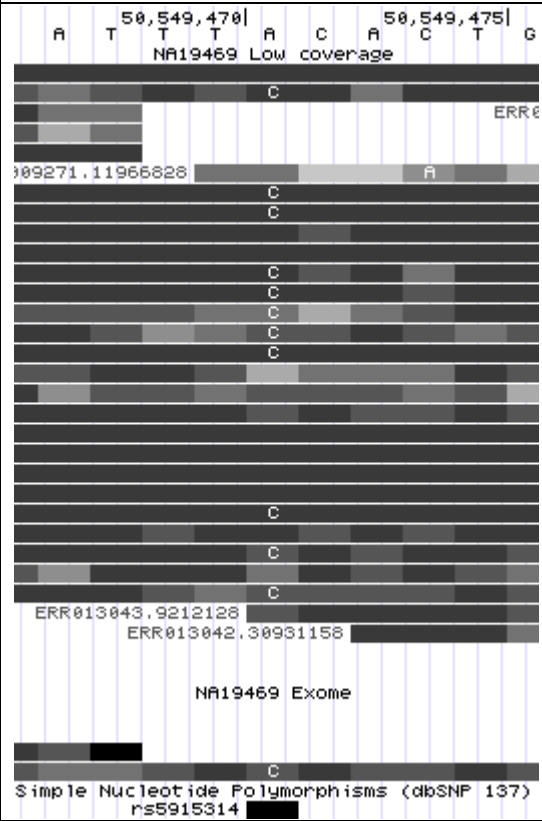
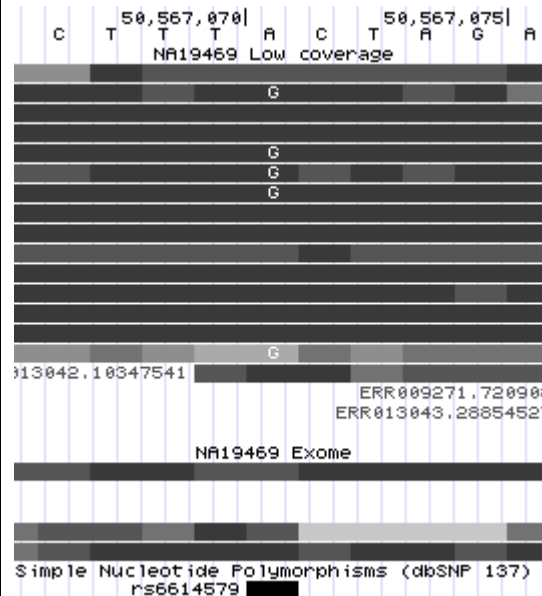
9・10 は VCF ファイルの genotype は正しくなさそうである。



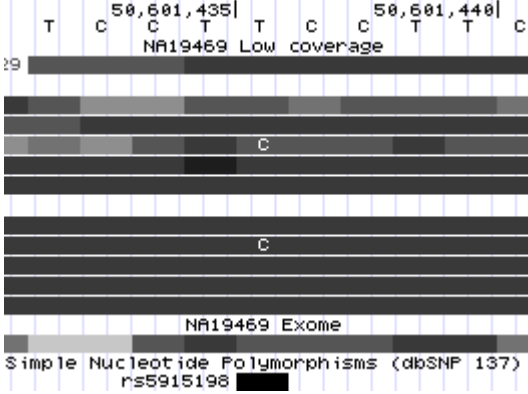
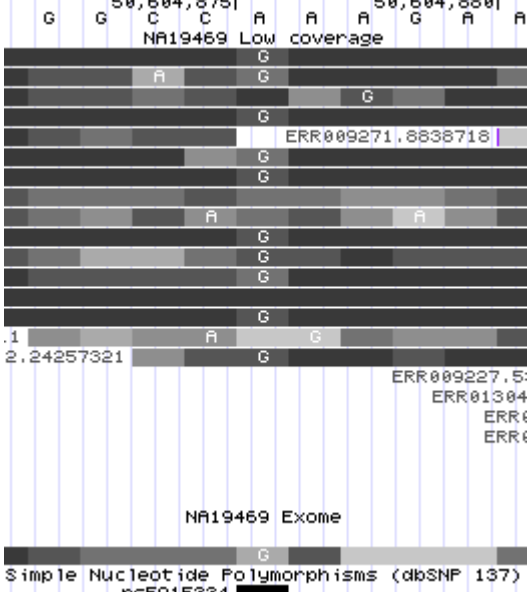
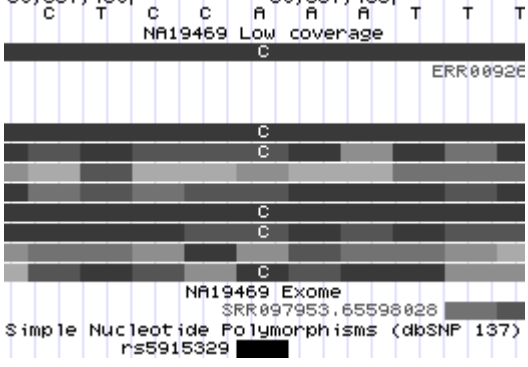
Base Quality minimum = 0

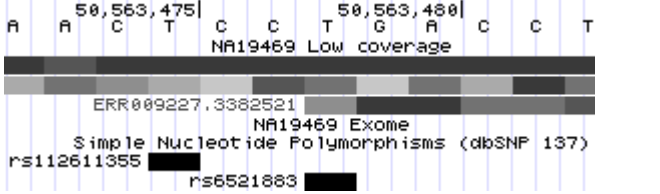
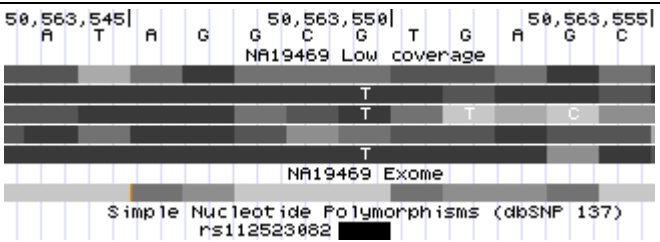
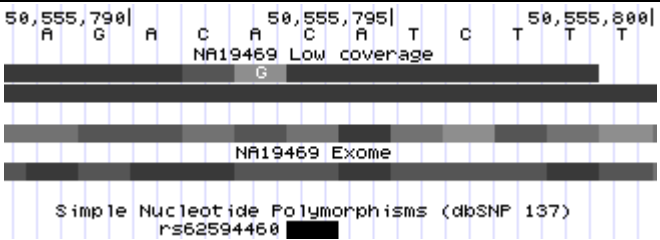
横軸: VCF が示す 2 つの allele の Quality value のうち大きい方

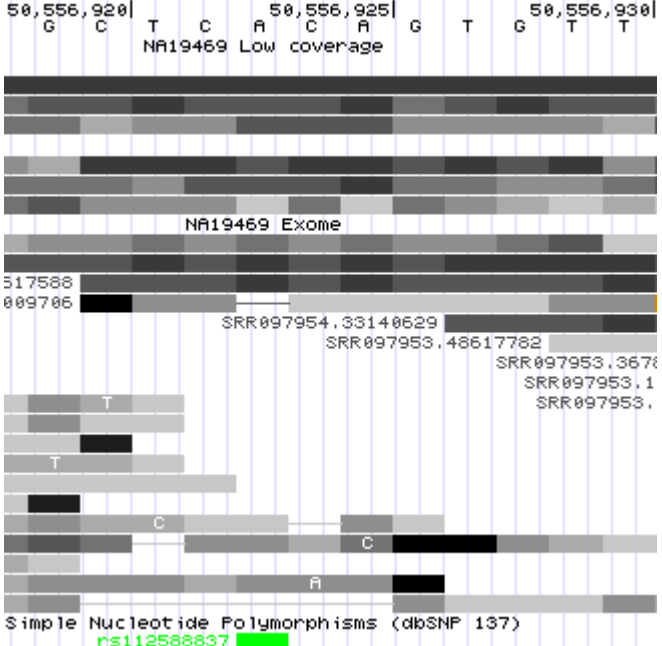
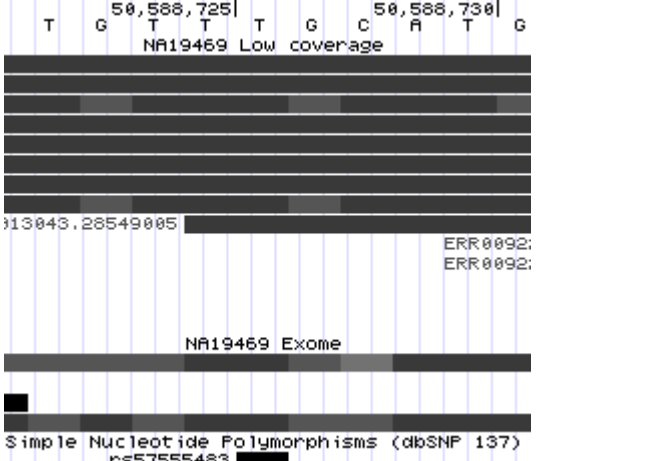
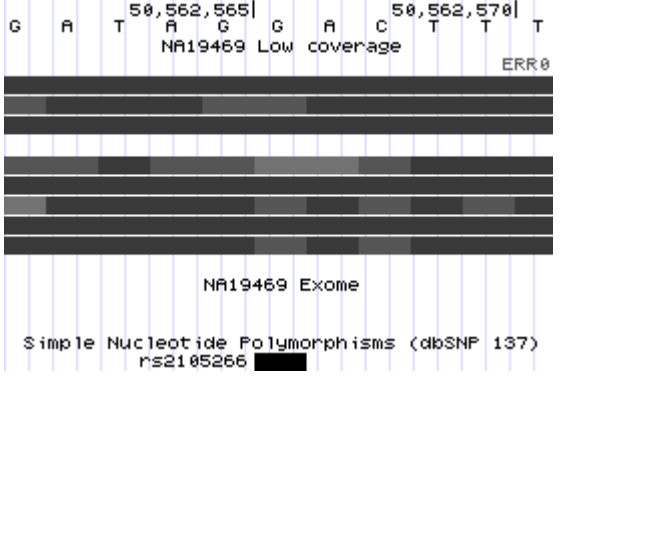
縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方

	UCSC Genome browser	
1		<p>chrX:50549471, REF=A, ALT=C</p> <p>VCF genotype = C   A</p> <p><math>Q_v(C)=161</math>, <math>Q_v(A)=147</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=14, 比=1.1</p> <p>正しい genotype: C   A (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=166</math>, <math>Q_v(A)=154</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=12, 比=1.1</p>
2		<p>chrX:50567071, REF=A, ALT=G</p> <p>VCF genotype = A   G</p> <p><math>Q_v(A)=218</math>, <math>Q_v(G)=96</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=122, 比=2.3</p> <p>正しい genotype: A   G (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(A)=218</math>, <math>Q_v(G)=99</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=119, 比=2.2</p>



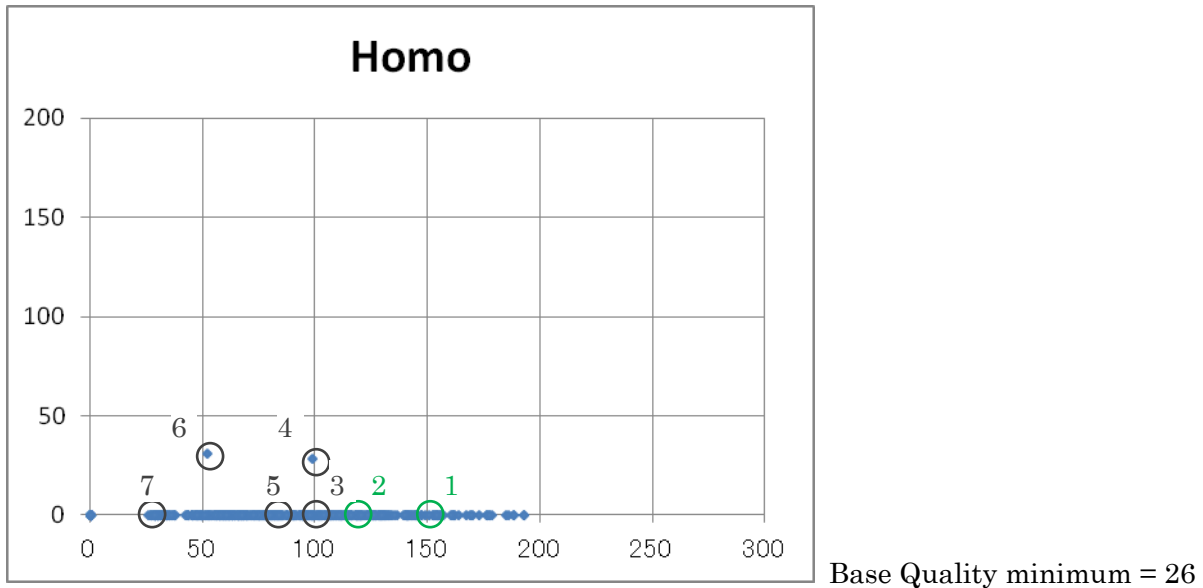
3		<p>chrX:50601436, REF=T, ALT=C</p> <p>VCF genotype = T   C</p> <p><math>Q_v(T)=157</math>, <math>Q_v(C)=61</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=96, 比=2.6</p> <p>正しい genotype: T   C (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(T)=165</math>, <math>Q_v(C)=61</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=104, 比=2.7</p>
4		<p>chrX:50604876, REF=A, ALT=G</p> <p>VCF genotype = G   A</p> <p><math>Q_v(A)=62</math>, <math>Q_v(G)=26</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=36, 比=2.4</p> <p>正しい genotype: G   A (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(G)=111</math>, <math>Q_v(A)=80</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=31, 比=1.4</p>
5		<p>chrX:50587483, REF=A, ALT=C</p> <p>VCF genotype = A   C</p> <p><math>Q_v(C)=100</math>, <math>Q_v(A)=27</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=73, 比=3.7</p> <p>正しい genotype: C/?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=100</math>, <math>Q_v(A)=50</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=50, 比=2.0</p>

6		<p>chrX:50563478, REF=T, ALT=C</p> <p>VCF genotype = T C</p> <p><math>Q_v(T)=32</math>, <math>Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=32, 比=×</p> <p>正しい genotype: T/?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(T)=55</math>, <math>Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=55, 比=×</p>
7		<p>chrX:50563550, REF=G, ALT=T</p> <p>VCF genotype = T G</p> <p><math>Q_v(T)=47</math>, <math>Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=47, 比=×</p> <p>正しい genotype: T/?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(T)=47</math>, <math>Q_v(G)=31</math>, <math>Q_v(A)=Q_v(C)=0</math></p> <p>差=16, 比=1.5</p>
8		<p>chrX:50555794, REF=C, ALT=T</p> <p>VCF genotype = T C</p> <p><math>Q_v(C)=74</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math></p> <p>差=74, 比=×</p> <p>正しい genotype: C/?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=96</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math></p> <p>差=96, 比=×</p>

<p>9</p>	 <p>50,556,920  50,556,925  50,556,930 </p> <p>T C A C A G T G T T</p> <p>NA19469 Low coverage</p> <p>NA19469 Exome</p> <p>517588</p> <p>809786</p> <p>SRR097954.33140629</p> <p>SRR097953.48617782</p> <p>SRR097953.3678</p> <p>SRR097953.1</p> <p>SRR097953.</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs112588837</p>	<p>chrX:50556923, REF=A, ALT=G</p> <p>VCF genotype = A   G</p> <p><math>Q_v(A)=99</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math></p> <p>差=99, 比=×</p> <p>正しい genotype: A/A</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(A)=150</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math></p> <p>差=150, 比=×</p>
<p>10</p>	 <p>50,588,725  50,588,730 </p> <p>T G T T T G C A T G</p> <p>NA19469 Low coverage</p> <p>NA19469 Exome</p> <p>313043.28549005</p> <p>ERR0092:</p> <p>ERR0092:</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs57555483</p>	<p>chrX:50588726, REF=T, ALT=C</p> <p>VCF genotype = T   C</p> <p><math>Q_v(T)=175</math>, <math>Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=175, 比=×</p> <p>正しい genotype: T/T</p> <p>Base Quality minimum = 0 のとき同じ</p>
<p>11</p>	 <p>50,562,565  50,562,570 </p> <p>G A T A G G A C T T T</p> <p>NA19469 Low coverage</p> <p>NA19469 Exome</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs2105266</p>	<p>chrX:50562566, REF=G, ALT=A</p> <p>VCF genotype = G   A</p> <p><math>Q_v(G)=113</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math></p> <p>差=113, 比=×</p> <p>正しい genotype: G/G</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(G)=117</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math></p> <p>差=117, 比=×</p>

② 個体名: HG00263

(a) VCF が Homo であるサイト

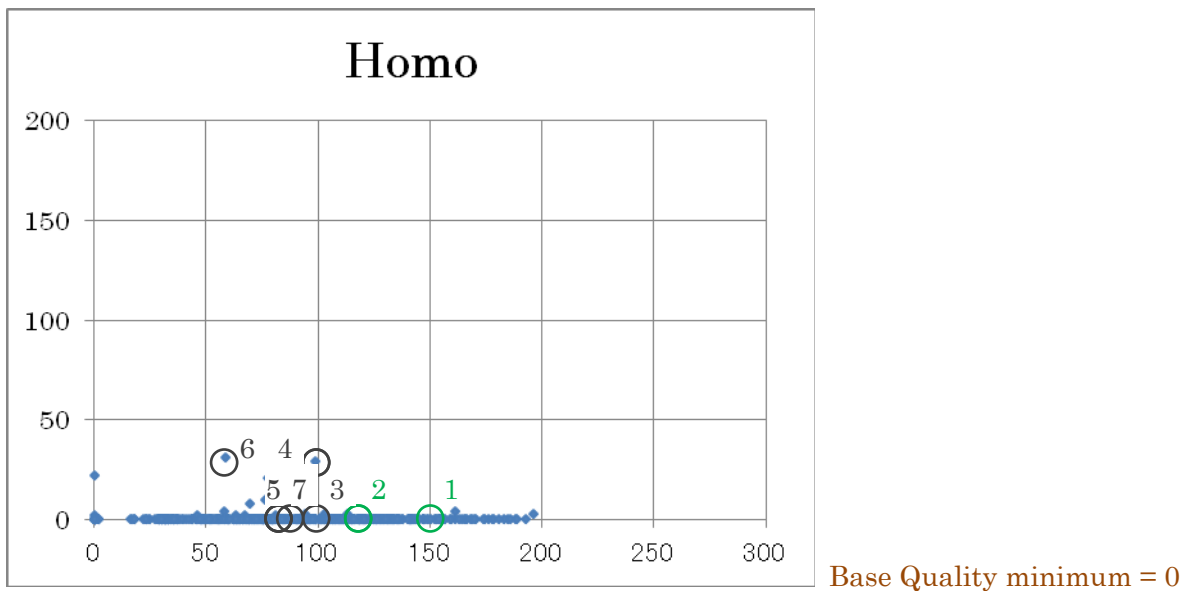


横軸: VCF が示す allele の Quality value

縦軸: それ以外の 3 つの allele の Quality value の最大値

1・2 は VCF ファイルの genotype が正しそうである。

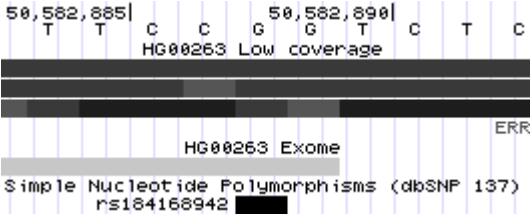
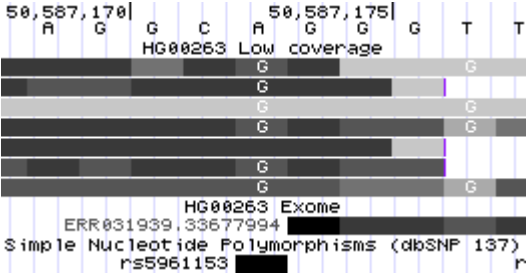
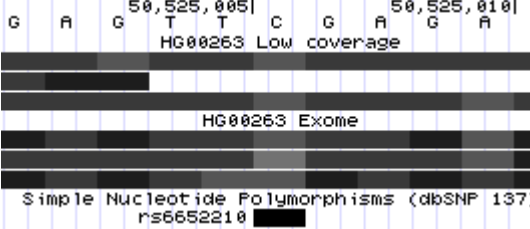
3～7 は VCF ファイルの genotype に矛盾は少ないが、リードの質が低いので判断が難しい。



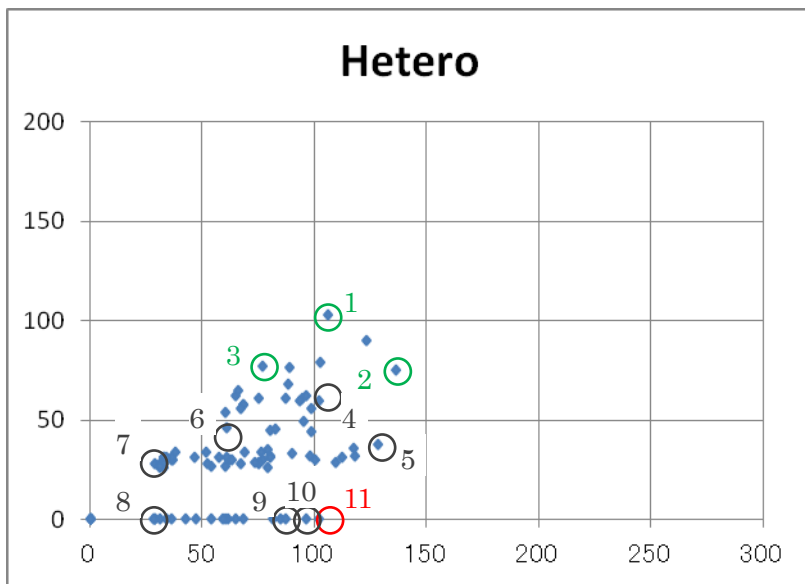
横軸: VCF が示す allele の Quality value

縦軸: それ以外の 3 つの allele の Quality value の最大値

UCSC Genome browser		
1	<p>chrX:50598090, REF=T, ALT=G VCF genotype = T   T Qv(T)=150, Qv(A)=Qv(G)=Qv(C)=0 差=150, 比=×</p> <p>正しい genotype: T   T (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>	
2	<p>chrX:50569963, REF=G, ALT=A VCF genotype = G   G Qv(G)=120, Qv(A)=Qv(T)=Qv(C)=0 差=120, 比=×</p> <p>正しい genotype: G   G (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>	
3	<p>chrX:50562576, REF=G, ALT=T VCF genotype = G   G Qv(G)=100, Qv(A)=Qv(T)=Qv(C)=0 差=100, 比=×</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき Qv(G)=101, Qv(A)=Qv(T)=Qv(C)=0 差=101, 比=×</p>	
4	<p>chrX:50597218, REF=C, ALT=T VCF genotype = C   C Qv(C)=99, Qv(T)=29, Qv(A)=Qv(G)=0 差=70, 比=3.4</p> <p>正しい genotype: C/?</p> <p>Base Quality minimum = 0 のとき同じ</p>	

5	 <p>50,582,885  C C G 50,582,890  C T C T T G G T HG00263 Low coverage HG00263 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs184168942</p>	<p>chrX:50582888, REF=G, ALT=A VCF genotype = G   G <math>Q_v(G)=80</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=80, 比=×</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(G)=82</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=82, 比=×</p>
6	 <p>50,587,170  G C A 50,587,175  G T T A G G G HG00263 Low coverage HG00263 Exome ERR031939.33677994 Simple Nucleotide Polymorphisms (dbSNP 137) rs5961153</p>	<p>chrX:50587173, REF=A, ALT=G VCF genotype = G   G <math>Q_v(G)=52</math>, <math>Q_v(A)=31</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=21, 比=1.7</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(G)=59</math>, <math>Q_v(A)=31</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=28, 比=1.9</p>
7	 <p>G A G 50,525,005  C G A 50,525,010  C T T T HG00263 Low coverage HG00263 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs6652210</p>	<p>chrX:50525006, REF=C, ALT=T VCF genotype = C   C <math>Q_v(C)=26</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=26, 比=×</p> <p>正しい genotype: C/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(C)=83</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=83, 比=×</p>

(b) VCF が Hetero であるサイト



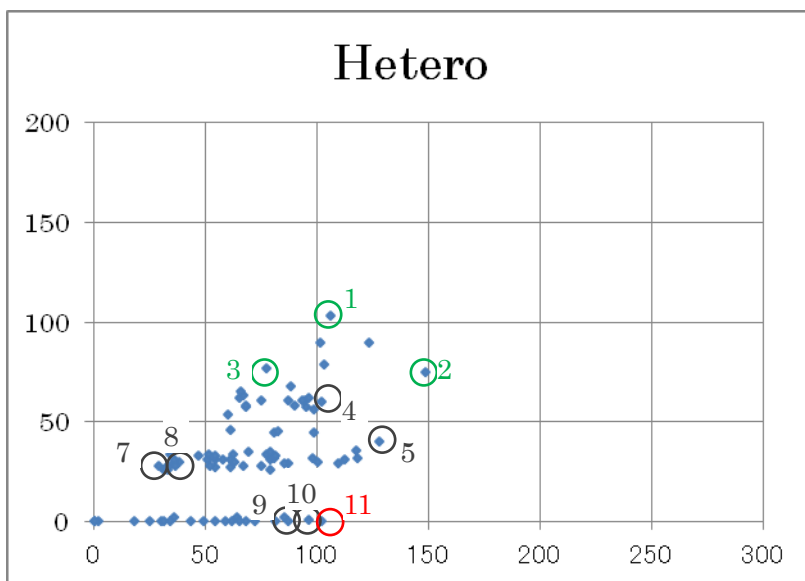
Base Quality minimum = 26

横軸: VCF が示す 2 つの allele の Quality value のうち大きい方  
縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方  
(VCF が示す allele 以外の残り 2 つはいずれも 0 であった。)

1～3 は VCF ファイルの genotype が正しそうである。

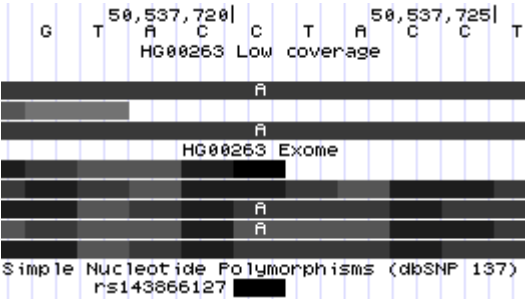
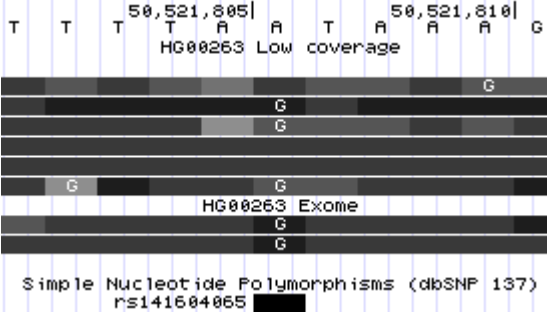
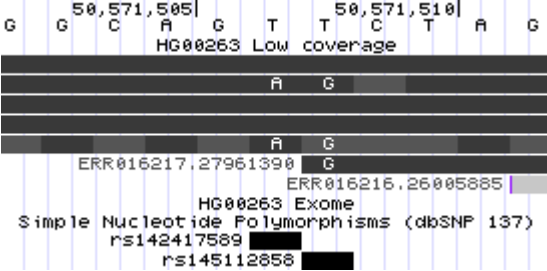
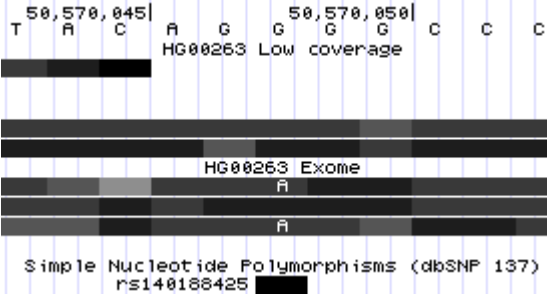
4～10 は VCF ファイルの genotype に矛盾は少ないが、リードの質が低いので判断が難しい。

11 は VCF ファイルの genotype は正しくなさそうである。

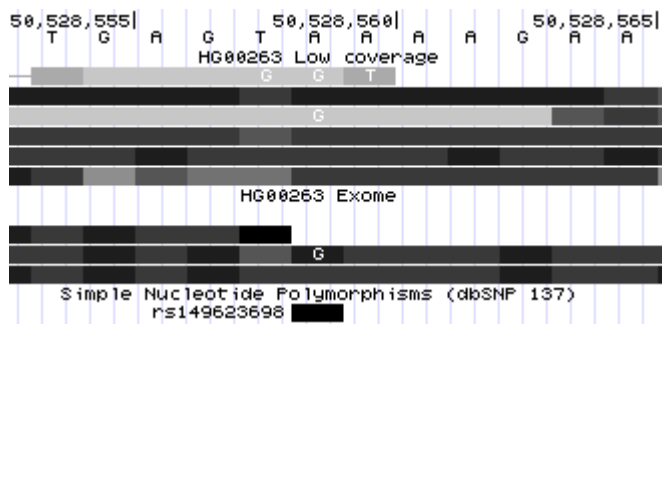
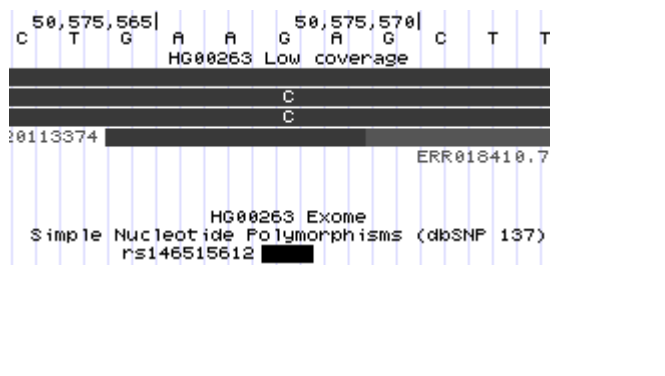
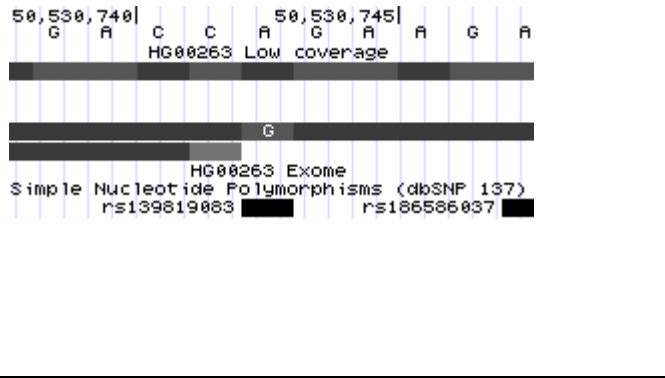
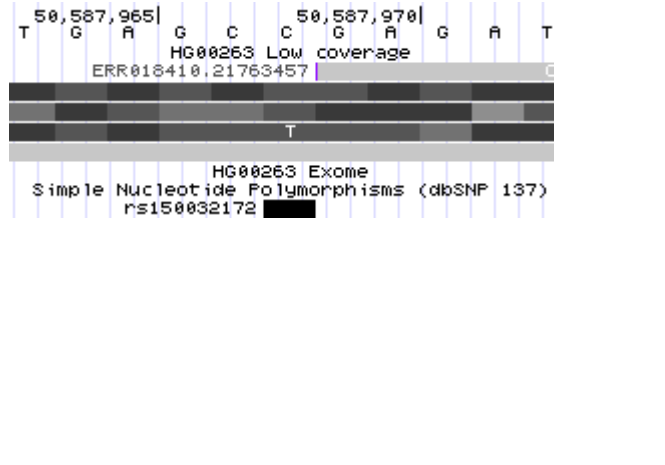


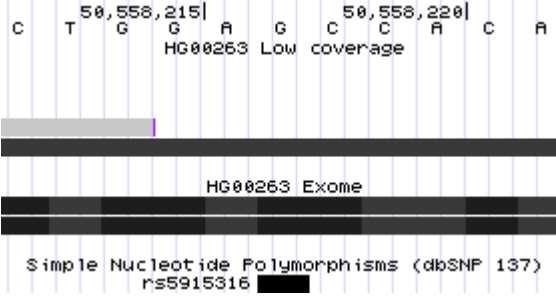
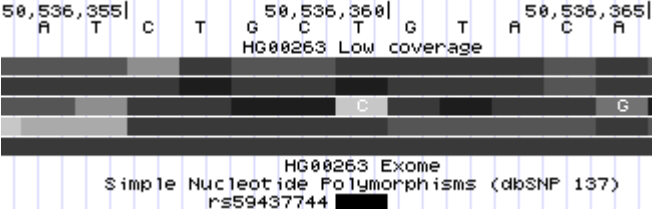
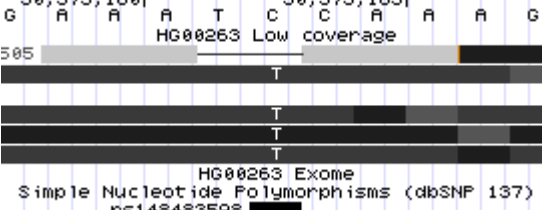
Base Quality minimum = 0

横軸: VCF が示す 2 つの allele の Quality value のうち大きい方  
縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方

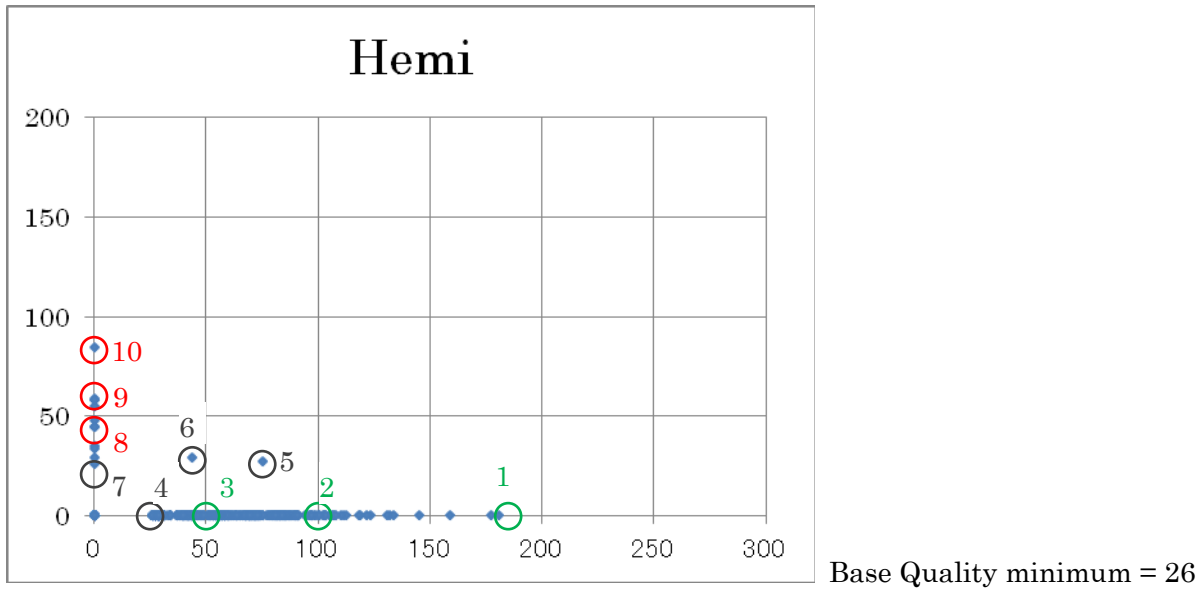
UCSC Genome browser		
1		<p>chrX:50537721, REF=C, ALT=A</p> <p>VCF genotype = A   C</p> <p><math>Q_v(A)=106</math>, <math>Q_v(C)=103</math>, <math>Q_v(G)=Q_v(T)=0</math></p> <p>差=3, 比=1.0</p> <p>正しい genotype: A   C (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>
2		<p>chrX:50521806, REF=A, ALT=G</p> <p>VCF genotype = G   A</p> <p><math>Q_v(G)=136</math>, <math>Q_v(A)=75</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=61, 比=1.8</p> <p>正しい genotype: G   A (=VCF)</p> <p>Base Quality minimum = 0 のとき  <math>Q_v(G)=149</math>, <math>Q_v(A)=75</math>, <math>Q_v(T)=Q_v(C)=0</math>          差=74, 比=2.0</p>
3		<p>chrX:50571508, REF=T, ALT=G</p> <p>VCF genotype = G   T</p> <p><math>Q_v(T)=Q_v(G)=77</math>, <math>Q_v(A)=Q_v(C)=0</math></p> <p>差=0, 比=1.0</p> <p>正しい genotype: G   T (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>
4		<p>chrX:50570048, REF=G, ALT=A</p> <p>VCF genotype = A   G</p> <p><math>Q_v(G)=102</math>, <math>Q_v(A)=60</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=42, 比=1.7</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき同じ</p>



5		<p>chrX:50528559, REF=A, ALT=G</p> <p>VCF genotype = G   A</p> <p><math>Q_v(A)=128</math>, <math>Q_v(G)=38</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=90, 比=3.4</p> <p>正しい genotype: A/?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(A)=128</math>, <math>Q_v(G)=40</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=88, 比=3.2</p>
6		<p>chrX:50575568, REF=G, ALT=C</p> <p>VCF genotype = C   G</p> <p><math>Q_v(G)=61</math>, <math>Q_v(C)=46</math>, <math>Q_v(A)=Q_v(T)=0</math></p> <p>差=15, 比=1.3</p> <p>正しい genotype: ?/?</p> <p>Base Quality minimum = 0 のとき同じ</p>
7		<p>chrX:50530743, REF=A, ALT=G</p> <p>VCF genotype = G   A</p> <p><math>Q_v(A)=29</math>, <math>Q_v(G)=28</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=1, 比=1.0</p> <p>正しい genotype: ?/?</p> <p>Base Quality minimum = 0 のとき同じ</p>
8		<p>chrX:50587968, REF=C, ALT=T</p> <p>VCF genotype = T   C</p> <p><math>Q_v(T)=28</math>, <math>Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=28, 比=×</p> <p>正しい genotype: ?/?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=37</math>, <math>Q_v(T)=28</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=9, 比=1.3</p>

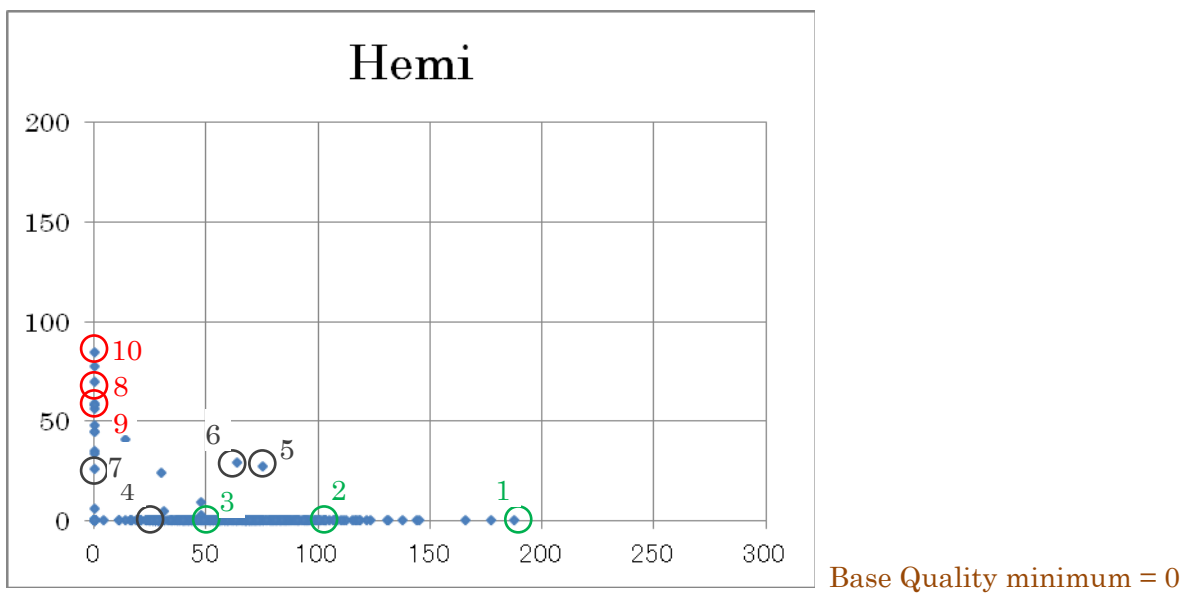
9	 <p>50,558,215  50,558,220  C T G G A G C C A C A HG00263 Low coverage</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137) rs5915316</p>	<p>chrX:50558217, REF=G, ALT=A</p> <p>VCF genotype = G   A</p> <p><math>Q_v(G)=87</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math></p> <p>差=87, 比=×</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき同じ</p>
10	 <p>50,536,355  50,536,360  50,536,365  A T C T G C T G T A C A HG00263 Low coverage</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137) rs59437744</p>	<p>chrX:50536360, REF=T, ALT=C</p> <p>VCF genotype = C   T</p> <p><math>Q_v(T)=96</math>, <math>Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=96, 比=×</p> <p>正しい genotype: T/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(T)=96</math>, <math>Q_v(C)=1</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=95, 比=96</p>
11	 <p>50,575,180  50,575,185  G A A A T C C A A A G HG00263 Low coverage</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137) rs148483598</p>	<p>chrX:50575183, REF=C, ALT=T</p> <p>VCF genotype = T   C</p> <p><math>Q_v(T)=102</math>, <math>Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=102, 比=×</p> <p>正しい genotype: T/T</p> <p>Base Quality minimum = 0 のとき同じ</p>

③ 個体名: NA20810



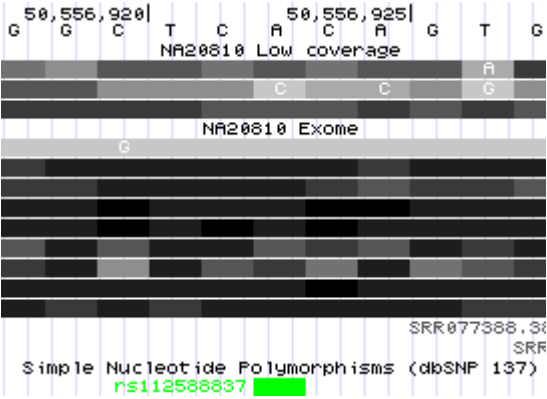
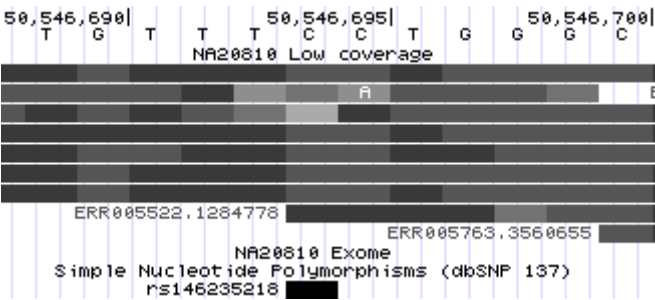
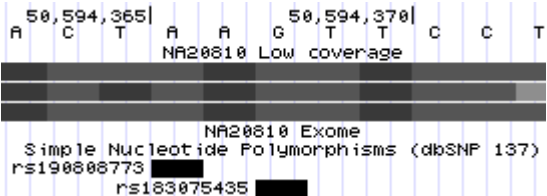
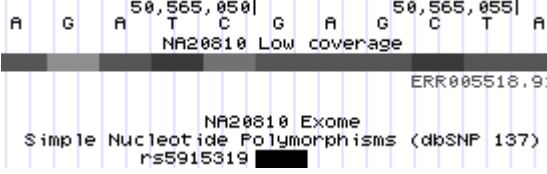
横軸: VCF が示す allele の Quality value

縦軸: それ以外の 3 つの allele の Quality value の最大値

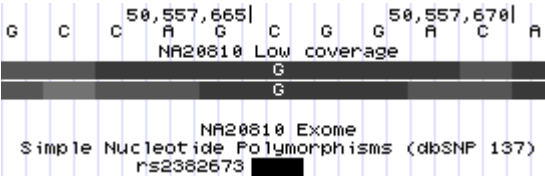
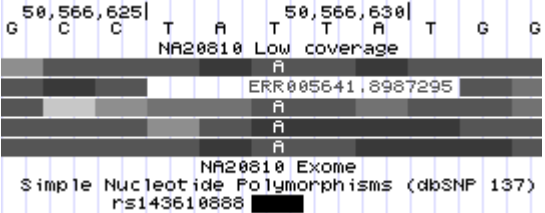


横軸: VCF が示す allele の Quality value

縦軸: それ以外の 3 つの allele の Quality value の最大値

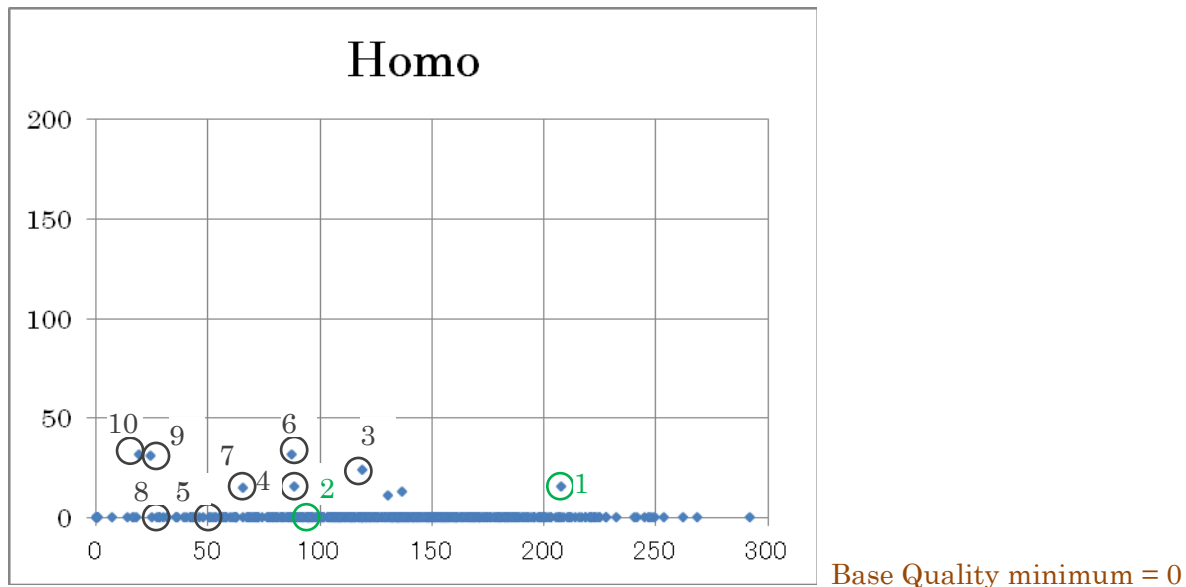
UCSC Gnome browser		
1	 <p>chrX:50556923, REF=A, ALT=G VCF genotype = A <math>Q_v(A)=181</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math> 差=181, 比=×</p> <p>正しい genotype: A (=VCF)</p> <p>Base Quality minimum = 0 のとき <math>Q_v(A)=188</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math> 差=188, 比=×</p>	
2	 <p>chrX:50546694, REF=C, ALT=T VCF genotype = C <math>Q_v(C)=100</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=100, 比=×</p> <p>正しい genotype: C (=VCF)</p> <p>Base Quality minimum = 0 のとき <math>Q_v(C)=107</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=107, 比=×</p>	
3	 <p>chrX:50594368, REF=G, ALT=T VCF genotype = G <math>Q_v(G)=51</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=51, 比=×</p> <p>正しい genotype: G (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>	
4	 <p>chrX:50565051, REF=G, ALT=A VCF genotype = G <math>Q_v(G)=26</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=26, 比=×</p> <p>正しい genotype: ?</p> <p>Base Quality minimum = 0 のとき同じ</p>	

5	<p>50,587,670  T A T C A A 50,587,675  C A G 50,587,680  A A</p> <p>NA20810 Low coverage</p> <p>ERR005519.6582491 G</p> <p>ERR005519.8374224</p> <p>NA20810 Exome</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs182104394</p>	<p>chrX:50587674, REF=A, ALT=G</p> <p>VCF genotype = A</p> <p><math>Q_v(A)=75, Q_v(G)=27, Q_v(T)=Q_v(C)=0</math></p> <p>差=48, 比=2.8</p> <p>正しい genotype: ?</p> <p>Base Quality minimum = 0 のとき同じ</p>
6	<p>T A 50,573,675  A A G 50,573,680  C A G G G</p> <p>NA20810 Low coverage</p> <p>C</p> <p>NA20810 Exome</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs72619080</p>	<p>chrX:50573677, REF=G, ALT=T</p> <p>VCF genotype = G</p> <p><math>Q_v(G)=44, Q_v(C)=29, Q_v(A)=Q_v(T)=0</math></p> <p>差=15, 比=1.5</p> <p>正しい genotype: ?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(G)=64, Q_v(C)=29, Q_v(A)=Q_v(T)=0</math></p> <p>差=35, 比=2.2</p>
7	<p>50,587,425  T A C T C 50,587,430  A A A 50,587,435  T A</p> <p>NA20810 Low coverage</p> <p>C C</p> <p>NA20810 Exome</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs5915328</p>	<p>chrX:50587430, REF=T, ALT=C</p> <p>VCF genotype = T</p> <p><math>Q_v(C)=26, Q_v(A)=Q_v(T)=Q_v(G)=0</math></p> <p>差=-26, 比=0</p> <p>正しい genotype: ?</p> <p>Base Quality minimum = 0 のとき同じ</p>
8	<p>50,546,470  T C T G T G 50,546,475  G T C T T 50,546,480  T A G</p> <p>NA20810 Low coverage</p> <p>G G G</p> <p>ERR005519</p> <p>NA20810 Exome</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs182458721</p> <p>rs138302819</p>	<p>chrX:50546475, REF=C, ALT=G</p> <p>VCF genotype = C</p> <p><math>Q_v(G)=45, Q_v(A)=Q_v(T)=Q_v(C)=0</math></p> <p>差=-45, 比=0</p> <p>正しい genotype: G</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(G)=70, Q_v(A)=Q_v(T)=Q_v(C)=0</math></p> <p>差=-70, 比=0</p>

9		<p>chrX:50557666, REF=C, ALT=G</p> <p>VCF genotype = C</p> <p>Qv(G)=59, Qv(A)=Qv(T)=Qv(C)=0</p> <p>差=-59, 比=0</p> <p>正しい genotype: G</p> <p>Base Quality minimum = 0 のとき同じ</p>
10		<p>chrX:50566628, REF=T, ALT=A</p> <p>VCF genotype = T</p> <p>Qv(A)=85, Qv(T)=Qv(G)=Qv(C)=0</p> <p>差=-85, 比=0</p> <p>正しい genotype: A</p> <p>Base Quality minimum = 0 のとき同じ</p>

④ 個体名: NA19036

(a) VCF が Homo であるサイト



横軸: VCF が示す allele の Quality value

縦軸: それ以外の 3 つの allele の Quality value の最大値

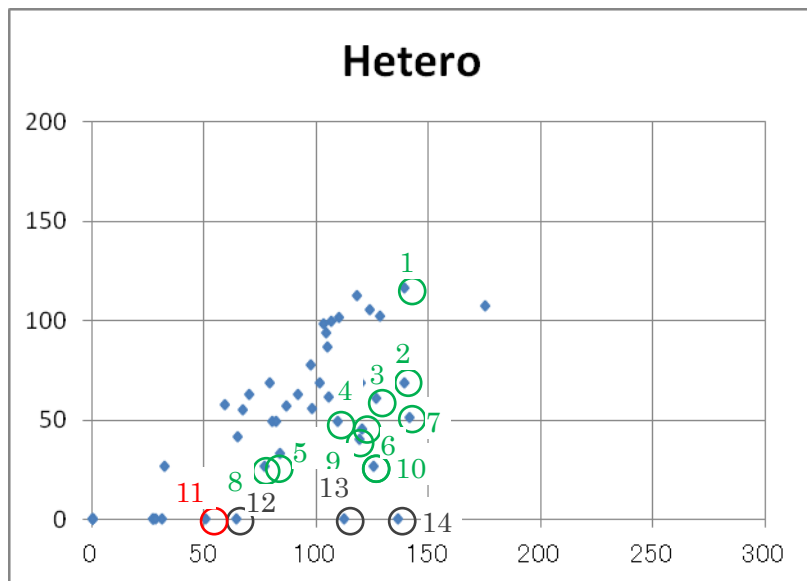
	UCSC Gnome browser	
1	<p>chrX:50577141, REF=C, ALT=T VCF genotype = T T Qv(T)=208, Qv(C)=16, Qv(A)=Qv(G)=0 差=192, 比=13 正しい genotype: T T (=VCF)</p>	
2	<p>chrX:50602945, REF=C, ALT=T VCF genotype = T T Qv(T)=95, Qv(A)=Qv(G)=Qv(C)=0 差=95, 比=× 正しい genotype: T T (=VCF)</p>	

3	<p>chrX:50552031, REF=T, ALT=C</p> <p>VCF genotype = T T</p> <p><math>Q_v(T)=119</math>, <math>Q_v(C)=24</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=95, 比=5.0</p> <p>正しい genotype: T/?</p>
4	<p>chrX:50602146, REF=C, ALT=T</p> <p>VCF genotype = C C</p> <p><math>Q_v(C)=89</math>, <math>Q_v(T)=16</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=73, 比=5.6</p> <p>正しい genotype: C/?</p>
5	<p>chrX:50563550, REF=G, ALT=T</p> <p>VCF genotype = G G</p> <p><math>Q_v(G)=51</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math></p> <p>差=51, 比=×</p> <p>正しい genotype: G/?</p>
6	<p>chrX:50526865, REF=C, ALT=T</p> <p>VCF genotype = C C</p> <p><math>Q_v(C)=87</math>, <math>Q_v(T)=32</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=55, 比=2.7</p> <p>正しい genotype: C/?</p>
7	<p>chrX:50560866, REF=T, ALT=C</p> <p>VCF genotype = T T</p> <p><math>Q_v(T)=66</math>, <math>Q_v(A)=15</math>, <math>Q_v(G)=Q_v(C)=0</math></p> <p>差=51, 比=4.4</p> <p>正しい genotype: T/?</p>



8	<p>chrX:50587968, REF=C, ALT=T VCF genotype = C   C <math>Q_v(C)=28</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=28, 比=×</p> <p>正しい genotype: C/?</p>	
9	<p>chrX:50589692, REF=C, ALT=T VCF genotype = C   C <math>Q_v(C)=24</math>, <math>Q_v(T)=31</math>, <math>Q_v(A)=Q_v(G)=0</math> 差=-7, 比=0.77</p> <p>正しい genotype: ?/?</p>	
10	<p>chrX:50589713, REF=G, ALT=A VCF genotype = G   G <math>Q_v(G)=19</math>, <math>Q_v(A)=32</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=-13, 比=0.59</p> <p>正しい genotype: ?/?</p>	

(b) VCF が Hetero であるサイト



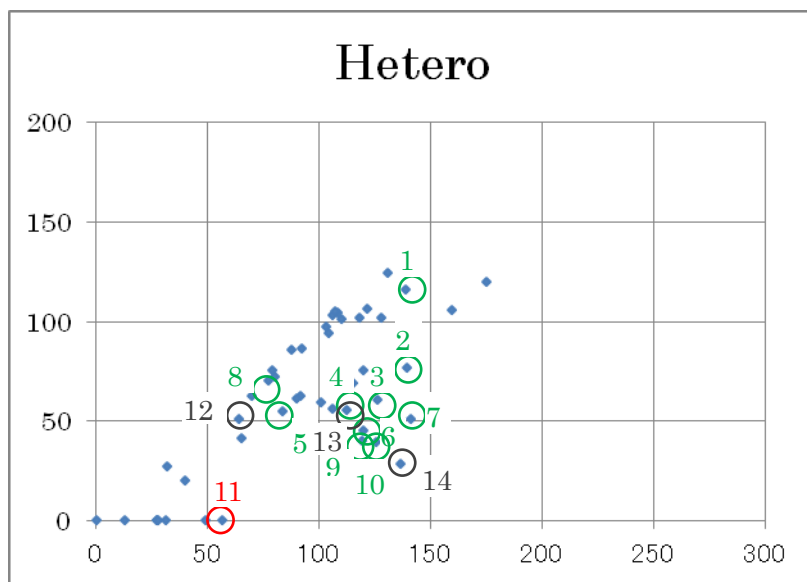
横軸: VCF が示す 2 つの allele の Quality value のうち大きい方

縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方

1～10 は VCF ファイルの genotype が正しそうである。(hetero)

12～14 は VCF ファイルの genotype に矛盾は少ないが、リードの質が低いので判断が難しい。

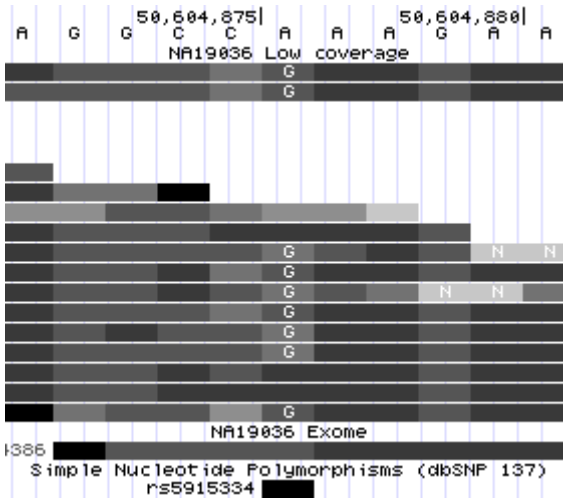
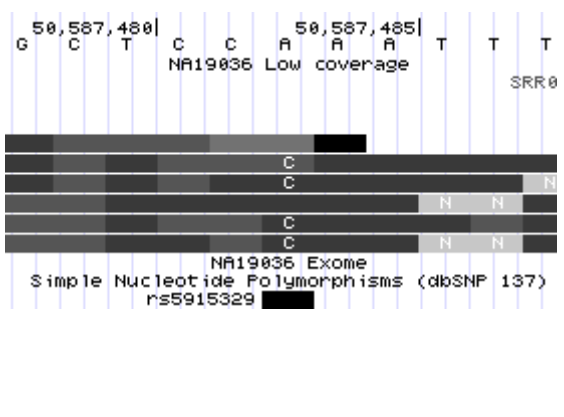
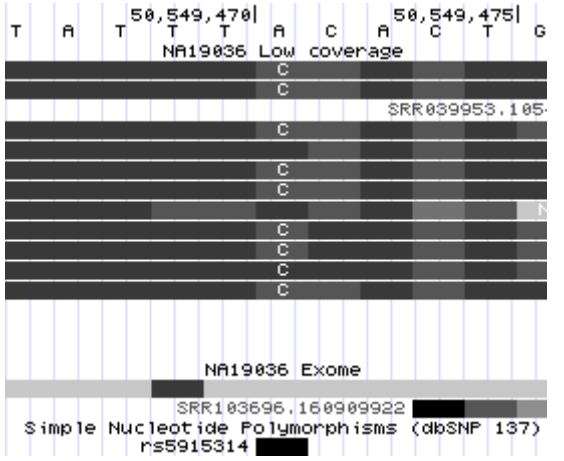
11 は VCF ファイルの genotype は正しくなさそうである。(homo)

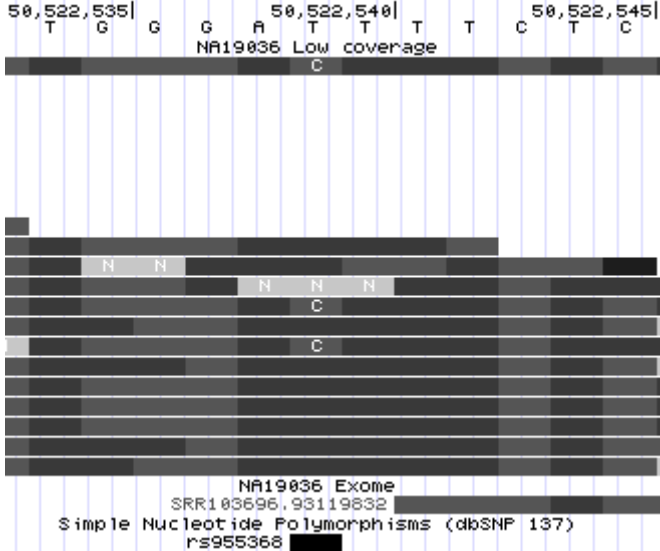
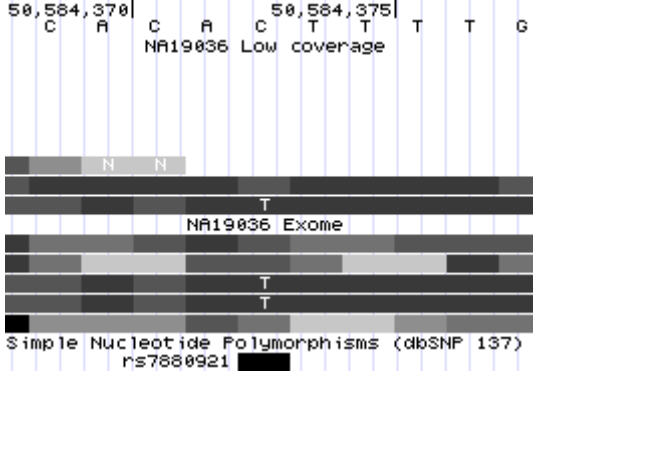
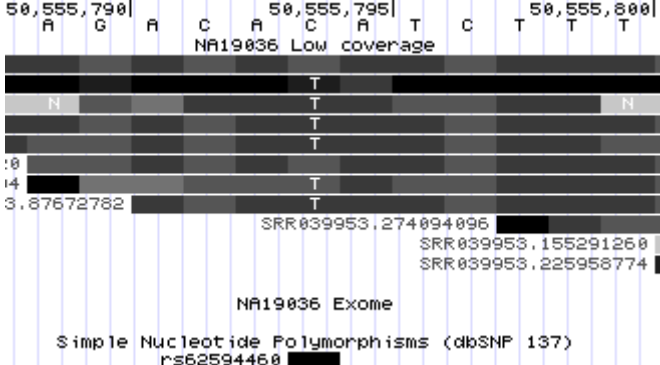
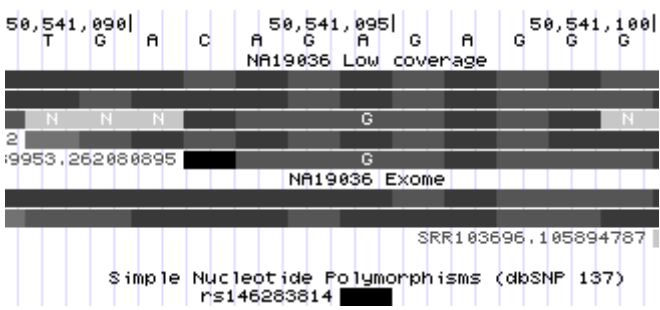


横軸: VCF が示す 2 つの allele の Quality value のうち大きい方

縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方

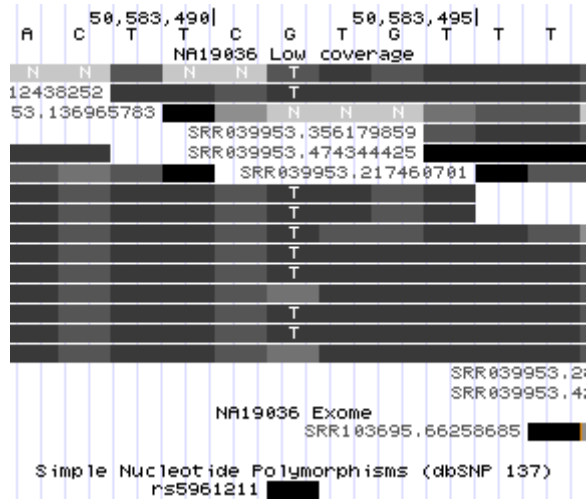
	UCSC Gnome browser	
1	<p>chrX:50578229, REF=G, ALT=T  VCF genotype = T G  <math>Q_v(T)=139</math>, <math>Q_v(G)=116</math>, <math>Q_v(A)=Q_v(C)=0</math>  差=23, 比=1.2</p> <p>正しい genotype: T G (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>	
2	<p>chrX:50602362, REF=G, ALT=A  VCF genotype = G A  <math>Q_v(A)=139</math>, <math>Q_v(G)=69</math>, <math>Q_v(T)=Q_v(C)=0</math>  差=70, 比=2.0</p> <p>正しい genotype: G A (=VCF)</p> <p>Base Quality minimum = 0 のとき  <math>Q_v(A)=139</math>, <math>Q_v(G)=77</math>, <math>Q_v(T)=Q_v(C)=0</math>  差=62, 比=1.8</p>	
3	<p>chrX:50567071, REF=A, ALT=G  VCF genotype = A G  <math>Q_v(G)=126</math>, <math>Q_v(A)=61</math>, <math>Q_v(T)=Q_v(C)=0</math>  差=65, 比=2.1</p> <p>正しい genotype: A G (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>	

4	 <p>chrX:50604876, REF=A, ALT=G</p> <p>VCF genotype = G   A</p> <p><math>Q_v(A)=110</math>, <math>Q_v(G)=50</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=60, 比=2.2</p> <p>正しい genotype: G   A (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(A)=118</math>, <math>Q_v(G)=102</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=16, 比=1.2</p>	<p>chrX:50604876, REF=A, ALT=G</p> <p>VCF genotype = G   A</p> <p><math>Q_v(A)=110</math>, <math>Q_v(G)=50</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=60, 比=2.2</p> <p>正しい genotype: G   A (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(A)=118</math>, <math>Q_v(G)=102</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=16, 比=1.2</p>
5	 <p>chrX:50587483, REF=A, ALT=C</p> <p>VCF genotype = C   A</p> <p><math>Q_v(C)=84</math>, <math>Q_v(A)=33</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=51, 比=2.5</p> <p>正しい genotype: C   A (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=84</math>, <math>Q_v(A)=55</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=29, 比=1.5</p>	<p>chrX:50587483, REF=A, ALT=C</p> <p>VCF genotype = C   A</p> <p><math>Q_v(C)=84</math>, <math>Q_v(A)=33</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=51, 比=2.5</p> <p>正しい genotype: C   A (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=84</math>, <math>Q_v(A)=55</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=29, 比=1.5</p>
6	 <p>chrX:50549471, REF=A, ALT=C</p> <p>VCF genotype = A   C</p> <p><math>Q_v(C)=120</math>, <math>Q_v(A)=46</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=74, 比=2.6</p> <p>正しい genotype: A   C (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=120</math>, <math>Q_v(A)=46</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=74, 比=2.6</p>	<p>chrX:50549471, REF=A, ALT=C</p> <p>VCF genotype = A   C</p> <p><math>Q_v(C)=120</math>, <math>Q_v(A)=46</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=74, 比=2.6</p> <p>正しい genotype: A   C (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(C)=120</math>, <math>Q_v(A)=46</math>, <math>Q_v(T)=Q_v(G)=0</math></p> <p>差=74, 比=2.6</p>

7	 <p>50,522,535  T G G 50,522,540  T T T C 50,522,545  T C</p> <p>NA19036 Low coverage</p> <p>NA19036 Exome</p> <p>SRR103696.93119832</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs955368</p>	<p>chrX:50522539, REF=T, ALT=C</p> <p>VCF genotype = T   C</p> <p><math>Q_v(T)=141</math>, <math>Q_v(C)=51</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=90, 比=2.8</p> <p>正しい genotype: T   C (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>
8	 <p>50,584,370  C A C 50,584,375  T T T G</p> <p>NA19036 Low coverage</p> <p>NA19036 Exome</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs7888921</p>	<p>chrX:50584373, REF=C, ALT=T</p> <p>VCF genotype = C   T</p> <p><math>Q_v(T)=77</math>, <math>Q_v(C)=27</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=50, 比=2.9</p> <p>正しい genotype: C   T (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(T)=77</math>, <math>Q_v(C)=70</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=7, 比=1.1</p>
9	 <p>50,555,790  A G A 50,555,795  C A T C 50,555,800  T T T</p> <p>NA19036 Low coverage</p> <p>NA19036 Exome</p> <p>SRR039953.274094096</p> <p>SRR039953.155291260</p> <p>SRR039953.225958774</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs62594460</p>	<p>chrX:50555794, REF=C, ALT=T</p> <p>VCF genotype = C   T</p> <p><math>Q_v(T)=119</math>, <math>Q_v(C)=41</math>, <math>Q_v(A)=Q_v(G)=0</math></p> <p>差=78, 比=2.9</p> <p>正しい genotype: C   T (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>
10	 <p>50,541,090  T G A C 50,541,095  G A G 50,541,100  G G G</p> <p>NA19036 Low coverage</p> <p>NA19036 Exome</p> <p>SRR103696.105894787</p> <p>Simple Nucleotide Polymorphisms (dbSNP 137)</p> <p>rs146283814</p>	<p>chrX:50541095, REF=A, ALT=G</p> <p>VCF genotype = A   G</p> <p><math>Q_v(A)=126</math>, <math>Q_v(G)=27</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=99, 比=4.7</p> <p>正しい genotype: A   G (=VCF)</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(A)=126</math>, <math>Q_v(G)=40</math>, <math>Q_v(T)=Q_v(C)=0</math></p> <p>差=86, 比=3.2</p>

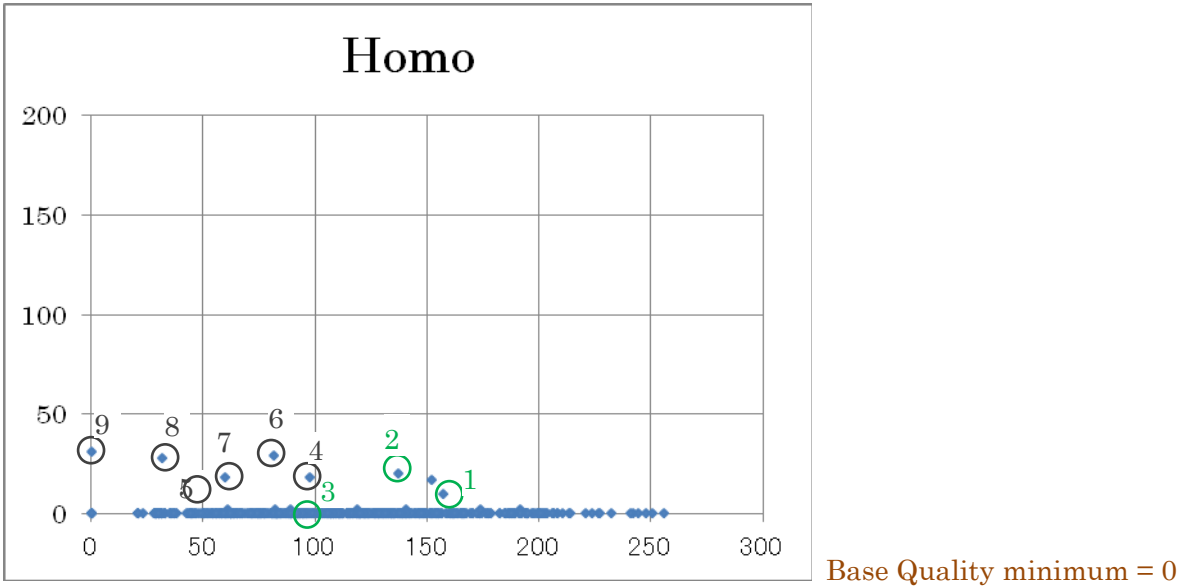
11	<p>chrX:50602941, REF=C, ALT=T  VCF genotype = C   T  <math>Q_v(C)=51</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math>  差=51, 比=×</p> <p>正しい genotype: C/C</p> <p>Base Quality minimum = 0 のとき  <math>Q_v(C)=57</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math>  差=57, 比=×</p>	
12	<p>chrX:50554716, REF=G, ALT=A  VCF genotype = G   A  <math>Q_v(A)=64</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math>  差=64, 比=×</p> <p>正しい genotype: A/?</p> <p>Base Quality minimum = 0 のとき  <math>Q_v(A)=64</math>, <math>Q_v(G)=51</math>, <math>Q_v(T)=Q_v(C)=0</math>  差=13, 比=1.3</p>	
13	<p>chrX:50557666, REF=C, ALT=G  VCF genotype = C   G  <math>Q_v(G)=113</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math>  差=113, 比=×</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき  <math>Q_v(G)=113</math>, <math>Q_v(C)=56</math>, <math>Q_v(A)=Q_v(T)=0</math>  差=57, 比=2.0</p>	

14



⑤ 個体名:HG00124

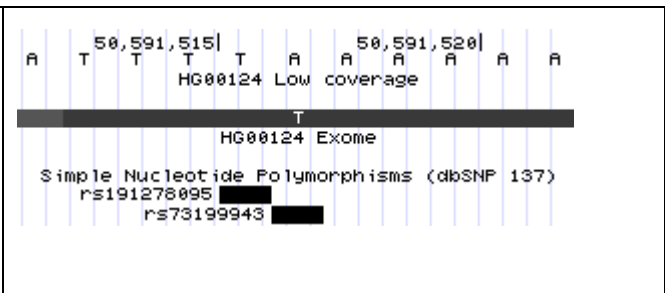
(a) VCF が Homo であるサイト



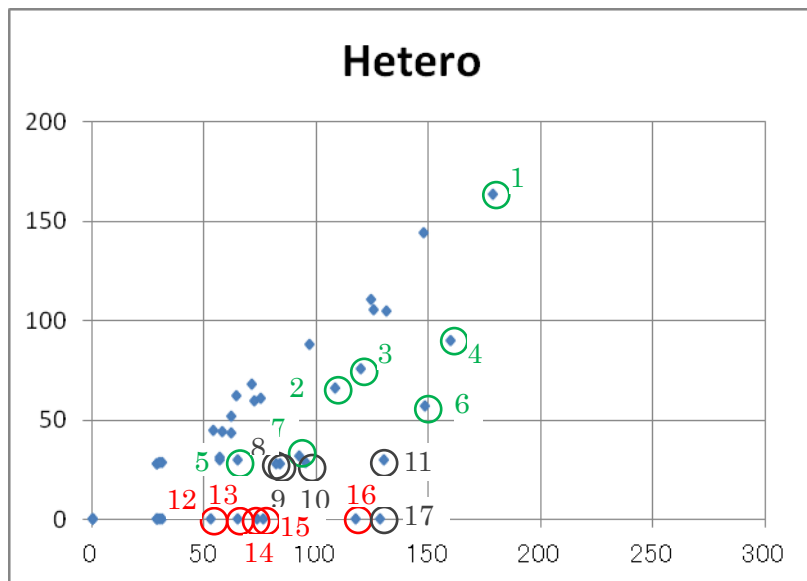
	UCSC Gnome browser	Base Quality minimum = 0
1		chrX:50581853, REF=C, ALT=T VCF genotype = C   C Qv(C)=157, Qv(A)=10, Qv(T)=Qv(G)=0 差=147, 比=16  正しい genotype: C   C (=VCF)
2		chrX:50547191, REF=C, ALT=T VCF genotype = C   C Qv(C)=137, Qv(A)=20, Qv(T)=Qv(G)=0 差=117, 比=6.9  正しい genotype: C   C (=VCF)



3	<p>chrX:50573760, REF=C, ALT=T VCF genotype = C   C <math>Q_v(C)=95, Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=95, 比=×</p> <p>正しい genotype: C   C (=VCF)</p>	
4	<p>chrX:50575580, REF=A, ALT=C VCF genotype = A   A <math>Q_v(A)=98, Q_v(G)=18, Q_v(T)=Q_v(C)=0</math> 差=80, 比=5.4</p> <p>正しい genotype: A/?</p>	
5	<p>chrX:50588333, REF=A, ALT=T VCF genotype = A   A <math>Q_v(A)=50, Q_v(T)=Q_v(G)=Q_v(C)=0</math> 差=50, 比=×</p> <p>正しい genotype: A/?</p>	
6	<p>chrX:50587181, REF=T, ALT=C VCF genotype = T   T <math>Q_v(T)=82, Q_v(C)=29, Q_v(A)=Q_v(G)=0</math> 差=53, 比=2.8</p> <p>正しい genotype: T/?</p>	
7	<p>chrX:50604915, REF=T, ALT=G VCF genotype = T   T <math>Q_v(T)=60, Q_v(C)=18, Q_v(A)=Q_v(G)=0</math> 差=42, 比=3.3</p> <p>正しい genotype: T/?</p>	
8	<p>chrX:50564989, REF=G, ALT=C VCF genotype = G   G <math>Q_v(G)=32, Q_v(A)=28, Q_v(T)=Q_v(C)=0</math> 差=4, 比=1.1</p> <p>正しい genotype: ??</p>	

9		<p>chrX:50591517, REF=A, ALT=T</p> <p>VCF genotype = A   A</p> <p><math>Q_v(T)=31, Q_v(A)=Q_v(G)=Q_v(C)=0</math></p> <p>差=-31, 比=0</p> <p>正しい genotype: ??</p>
---	---	--

(b) VCF が Hetero であるサイト



Base Quality minimum = 26

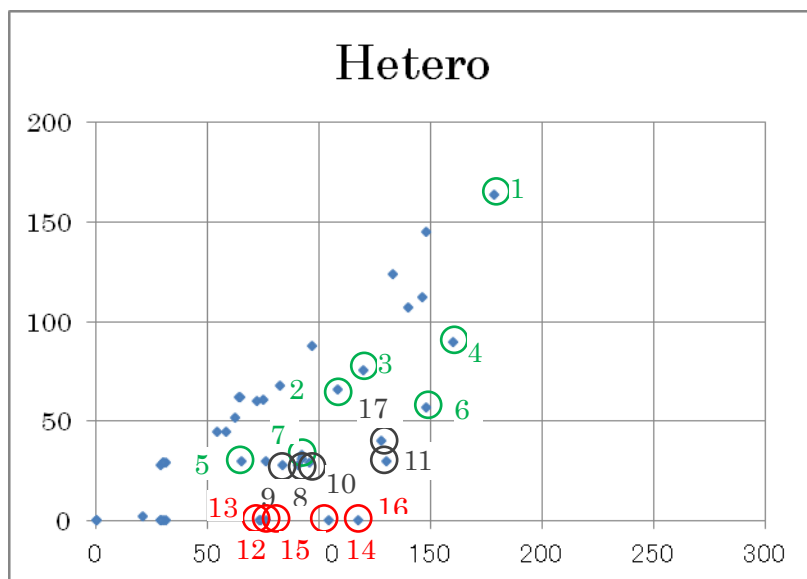
横軸: VCF が示す 2 つの allele の Quality value のうち大きい方

縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方

1～7 は VCF ファイルの genotype が正しそうである。(hetero)

8～11・17 は VCF ファイルの genotype に矛盾は少ないが、リードの質が低いので判断が難しい。

12～16 は VCF ファイルの genotype は正しくなさそうである。(homo)



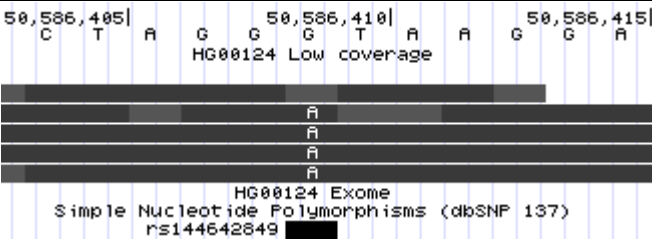
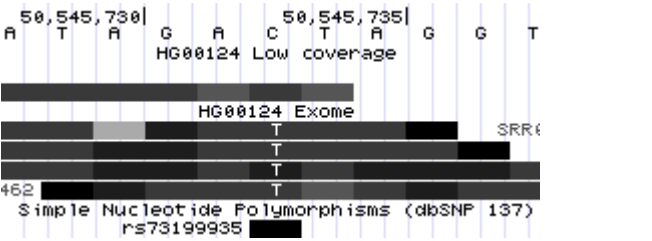
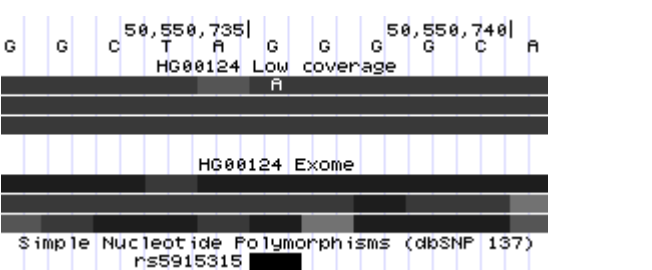
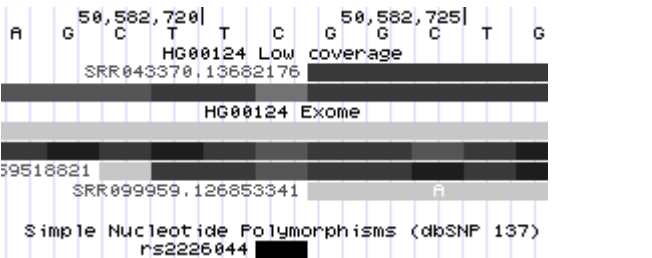
Ba0se Quality minimum = 0

横軸: VCF が示す 2 つの allele の Quality value のうち大きい方

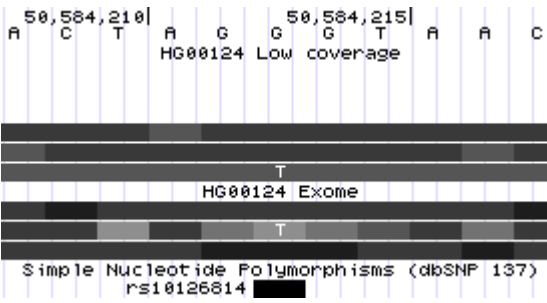
縦軸: VCF が示す 2 つの allele の Quality value のうち小さい方

	UCSC Gnome browser	
1	<p>chrX:50542692, REF=G, ALT=A VCF genotype = G   A Qv(G)=179, Qv(A)=164, Qv(T)=Qv(C)=0 差=15, 比=1.1 正しい genotype: G   A (=VCF) Base Quality minimum = 0 のとき同じ</p>	
2	<p>chrX:50602941, REF=C, ALT=T VCF genotype = T   C Qv(C)=109, Qv(T)=66, Qv(A)=Qv(G)=0 差=43, 比=1.7 正しい genotype: T   C (=VCF) Base Quality minimum = 0 のとき同じ</p>	
3	<p>chrX:50603361, REF=A, ALT=G VCF genotype = A   G Qv(G)=120, Qv(A)=76, Qv(T)=Qv(C)=0 差=44, 比=1.6 正しい genotype: A   G (=VCF) Base Quality minimum = 0 のとき同じ</p>	
4	<p>chrX:50547506, REF=C, ALT=T VCF genotype = C   T Qv(C)=160, Qv(T)=90, Qv(A)=Qv(G)=0 差=70, 比=1.8 正しい genotype: C   T (=VCF) Base Quality minimum = 0 のとき同じ</p>	

5	<p>chrX:50587483, REF=A, ALT=C VCF genotype = A   C <math>Q_v(A)=65</math>, <math>Q_v(C)=30</math>, <math>Q_v(T)=Q_v(G)=0</math> 差=35, 比=2.2</p> <p>正しい genotype: A   C (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>	
6	<p>chrX:50567071, REF=A, ALT=G VCF genotype = A   G <math>Q_v(G)=148</math>, <math>Q_v(A)=57</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=91, 比=2.6</p> <p>正しい genotype: A   G (=VCF)</p> <p>Base Quality minimum = 0 のとき同じ</p>	
7	<p>chrX:50549471, REF=A, ALT=C VCF genotype = A   C <math>Q_v(C)=92</math>, <math>Q_v(A)=32</math>, <math>Q_v(T)=Q_v(G)=0</math> 差=60, 比=2.9</p> <p>正しい genotype: A   C (=VCF)</p> <p>Base Quality minimum = 0 のとき <math>Q_v(C)=92</math>, <math>Q_v(A)=33</math>, <math>Q_v(T)=Q_v(G)=0</math> 差=59, 比=2.8</p>	
8	<p>chrX:50604876, REF=A, ALT=G VCF genotype = A   G <math>Q_v(A)=82</math>, <math>Q_v(G)=28</math>, <math>Q_v(T)=Q_v(C)=0</math> 差=54, 比=2.9</p> <p>正しい genotype: A/?</p> <p>Base Quality minimum = 0 のとき <math>Q_v(A)=90</math>, <math>Q_v(T)=28</math>, <math>Q_v(G)=Q_v(C)=0</math> 差=62, 比=3.2</p>	

9	 <p>50,586,405  50,586,410  50,586,415  C T A G G G T A A G G A HG00124 Low coverage HG00124 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs144642849</p>	<p>chrX:50586409, REF=G, ALT=A VCF genotype = G A Qv(A)=84, Qv(G)=28, Qv(T)=Qv(C)=0 差=56, 比=3.0</p> <p>正しい genotype: A/?</p> <p>Base Quality minimum = 0 のとき同じ</p>
10	 <p>50,545,730  50,545,735  A T A G A C T A G G T HG00124 Low coverage HG00124 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs73199935</p>	<p>chrX:50545733, REF=C, ALT=T VCF genotype = T C Qv(T)=95, Qv(C)=29, Qv(A)=Qv(G)=0 差=66, 比=3.3</p> <p>正しい genotype: T/?</p> <p>Base Quality minimum = 0 のとき同じ</p>
11	 <p>50,550,735  50,550,740  G G C T A G G G G C A HG00124 Low coverage HG00124 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs5915315</p>	<p>chrX:50550736, REF=G, ALT=A VCF genotype = G A Qv(G)=130, Qv(A)=30, Qv(T)=Qv(C)=0 差=100, 比=4.3</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき同じ</p>
12	 <p>50,582,720  50,582,725  A G C T C G G C T G HG00124 Low coverage HG00124 Exome Simple Nucleotide Polymorphisms (dbSNP 137) rs2226044</p>	<p>chrX:50582722, REF=C, ALT=T VCF genotype = T C Qv(C)=53, Qv(A)=Qv(T)=Qv(G)=0 差=53, 比=×</p> <p>正しい genotype: C/C</p> <p>Base Quality minimum = 0 のとき Qv(C)=74, Qv(A)=Qv(T)=Qv(G)=0 差=74, 比=×</p>

13	<p>chrX:50578229, REF=G, ALT=T VCF genotype = T G <math>Q_v(G)=65</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=65, 比=×</p> <p>正しそうな genotype: G/G</p> <p>Base Quality minimum = 0 のとき <math>Q_v(G)=73</math>, <math>Q_v(A)=Q_v(T)=Q_v(C)=0</math> 差=73, 比=×</p>	
14	<p>chrX:50602971, REF=G, ALT=A VCF genotype = G A <math>Q_v(A)=74</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math> 差=74, 比=×</p> <p>正しそうな genotype: A/A</p> <p>Base Quality minimum = 0 のとき <math>Q_v(A)=105</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math> 差=105, 比=×</p>	
15	<p>chrX:50578254, REF=T, ALT=A VCF genotype = T A <math>Q_v(A)=76</math>, <math>Q_v(T)=Q_v(G)=Q_v(C)=0</math> 差=76, 比=×</p> <p>正しそうな genotype: A/A</p> <p>Base Quality minimum = 0 のとき同じ</p>	
16	<p>chrX:50601436, REF=T, ALT=C VCF genotype = C T <math>Q_v(C)=118</math>, <math>Q_v(A)=Q_v(T)=Q_v(G)=0</math> 差=118, 比=×</p> <p>正しそうな genotype: C/C</p> <p>Base Quality minimum = 0 のとき同じ</p>	

17		<p>chrX:50584213, REF=G, ALT=T</p> <p>VCF genotype = G   T</p> <p><math>Q_v(G)=128, Q_v(A)=Q_v(T)=Q_v(C)=0</math></p> <p>差=128, 比=×</p> <p>正しい genotype: G/?</p> <p>Base Quality minimum = 0 のとき</p> <p><math>Q_v(G)=128, Q_v(T)=40, Q_v(A)=Q_v(C)=0</math></p> <p>差=88, 比=3.2</p>
----	---	--