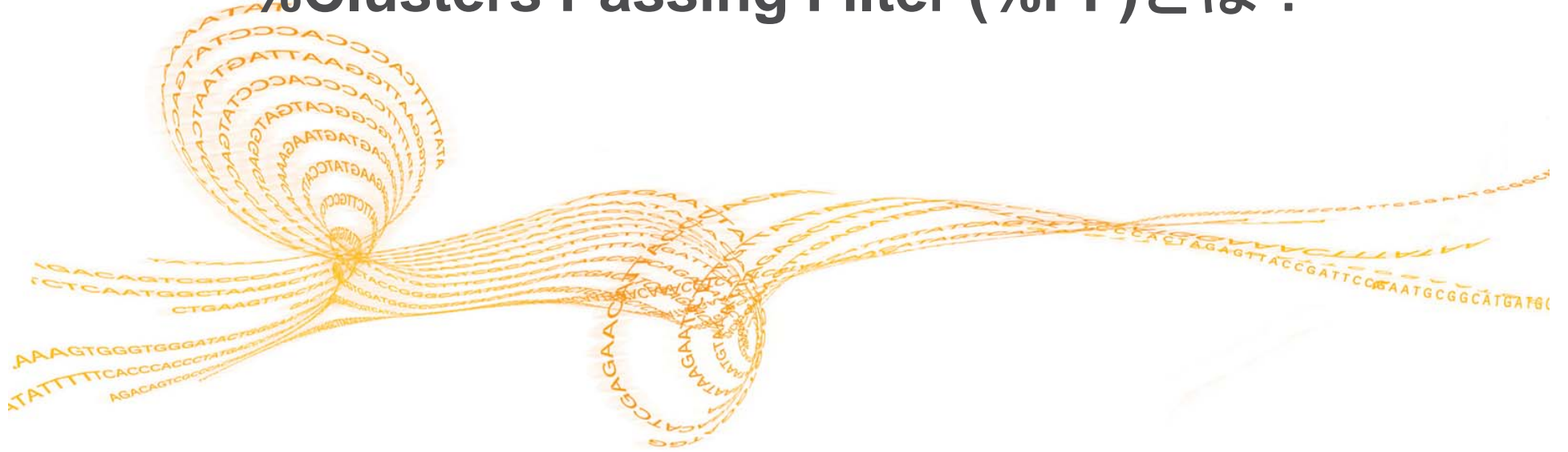
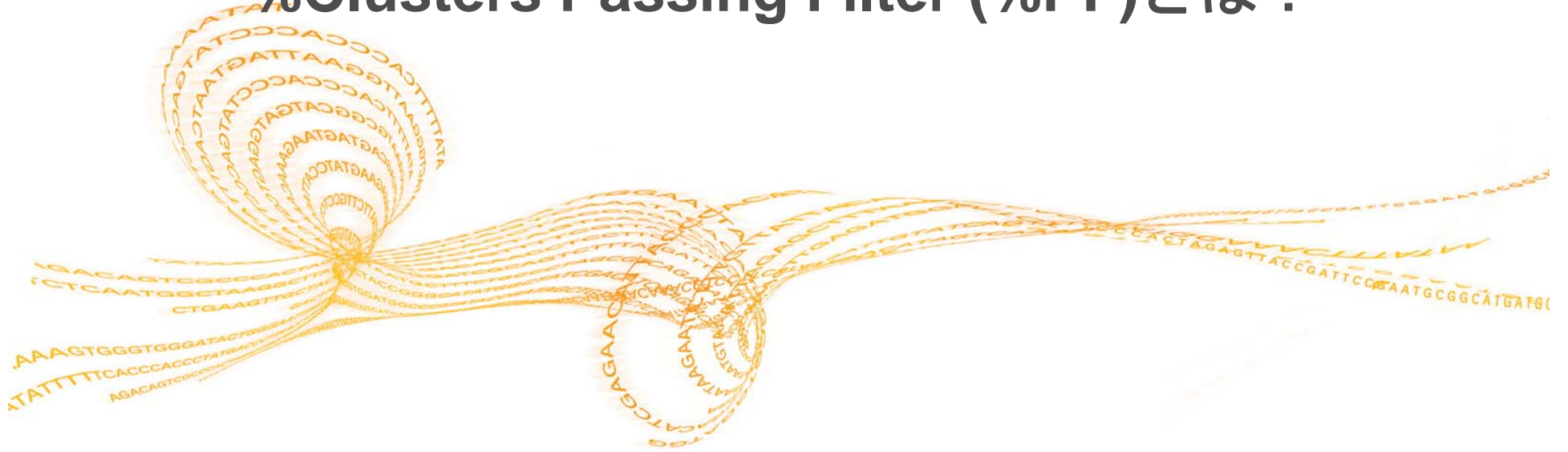


%Clusters Passing Filter (%PF)とは？



- ▶ %Clusters Passing Filter (%PF)の定義
- ▶ どのくらいの %PFが得られるのか
- ▶ %PFが低くなる主な原因とその診断方法
- ▶ %PFを最適化するには

%Clusters Passing Filter (%PF)とは？



- ▶ **%Clusters Passing Filter (%PF)の定義**
- ▶ どのくらいの %PFが得られるのか
- ▶ %PFが低くなる主な原因とその診断方法
- ▶ %PFを最適化するには

%Clusters Passing Filter (%PF)とは？

$$\%PF = \frac{\# \text{ clusters PF}}{\text{Density (k/mm}^2\text{)}} \times 100$$

Read 1			
Lane	Tiles	Density (K/mm ²)	Cluster PF (%)
1	96	530 +/- 71	95.09 +/- 1.12
2	96	734 +/- 89	91.90 +/- 1.74
3	96	135 +/- 131	12.38 +/- 23.85
4	96	44 +/- 5	0.00 +/- 0.00
5	96	228 +/- 32	89.58 +/- 1.37
6	96	399 +/- 28	69.94 +/- 21.36
7	96	352 +/- 36	82.49 +/- 2.33
8	96	39 +/- 8	0.00 +/- 0.00

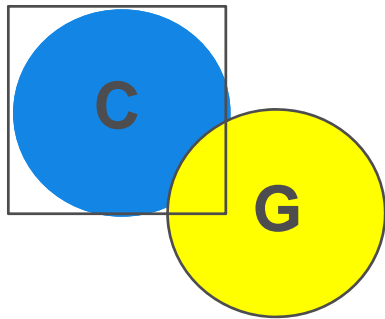
Screenshot from SAV Summary tab

5
 GTATCATAGATTAATTGATCCACTGATTCAGCTTACCTTAACGGANDGTATCAATDAGACTAAATATAACGTATCATTAAGAGCTACCGTCTCCGGCCGGAAGAAATGATAAAGATTAACACACTTCTGTTAACCTTAAGCGAAGCTATCATTAAGATTA
 ACTTAACACACTTGTGTTAAGCTTAAAGATTAATTTGATCCACTGATTCAGCTTACCTTAACGGANDGTATCAATDAGACTAAATATAACGTATCATTAAGAGCTACCGTCTCCGGCCGGAAGAAATGATAAAGATTAACACACTTCTGTTAACCTTAAGCGAAGCTATCATTAAGATTA
 CTTGCAACACTTAAAGACTTGTGTTAAGCTTAAAGATTAATTTGATCCACTGATTCAGCTTACCTTAACGGANDGTATCAATDAGACTAAATATAACGTATCATTAAGAGCTACCGTCTCCGGCCGGAAGAAATGATAAAGATTAACACACTTCTGTTAACCTTAAGCGAAGCTATCATTAAGATTA
 CTCTCTAAACCTTAAAGACTTGTGTTAAGCTTAAAGATTAATTTGATCCACTGATTCAGCTTACCTTAACGGANDGTATCAATDAGACTAAATATAACGTATCATTAAGAGCTACCGTCTCCGGCCGGAAGAAATGATAAAGATTAACACACTTCTGTTAACCTTAAGCGAAGCTATCATTAAGATTA
 CAACCTTAAAGACTTGTGTTAAGCTTAAAGATTAATTTGATCCACTGATTCAGCTTACCTTAACGGANDGTATCAATDAGACTAAATATAACGTATCATTAAGAGCTACCGTCTCCGGCCGGAAGAAATGATAAAGATTAACACACTTCTGTTAACCTTAAGCGAAGCTATCATTAAGATTA
 ATTAACCTTAAAGACTTGTGTTAAGCTTAAAGATTAATTTGATCCACTGATTCAGCTTACCTTAACGGANDGTATCAATDAGACTAAATATAACGTATCATTAAGAGCTACCGTCTCCGGCCGGAAGAAATGATAAAGATTAACACACTTCTGTTAACCTTAAGCGAAGCTATCATTAAGATTA
 BTATCATTAAGACTTGTGTTAAGCTTAAAGATTAATTTGATCCACTGATTCAGCTTACCTTAACGGANDGTATCAATDAGACTAAATATAACGTATCATTAAGAGCTACCGTCTCCGGCCGGAAGAAATGATAAAGATTAACACACTTCTGTTAACCTTAAGCGAAGCTATCATTAAGATTA

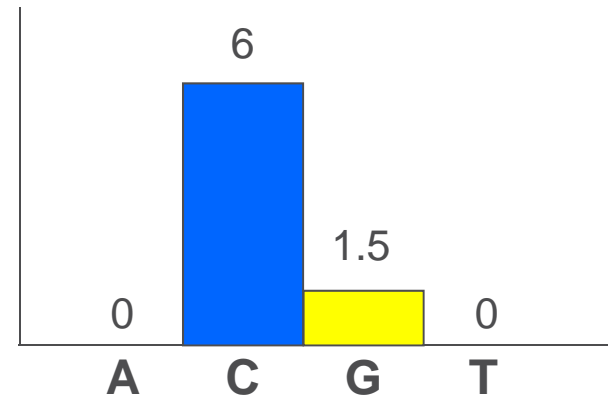
Quality Filter

The Chastity Filter

Cluster signal



重なるClusterからのシグナルを捨てるが、Chastityは0.8でありまだ当該クラスタのBase Callに影響を及ぼさない



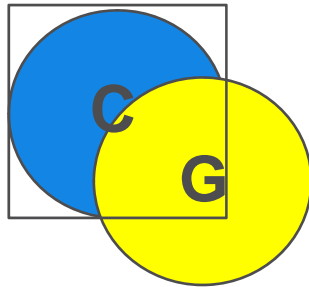
$$6 / (6 + 1.5) = 0.8$$

Chastityの計算式

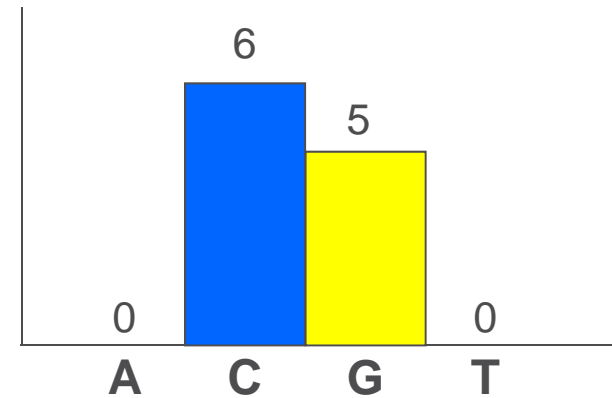
Quality Filter

The Chastity Filter

Cluster signal



十分にクラスターが重なってしまうと、クラスターのシグナルがどの塩基由来であるか判別できない。
そこで25サイクルまでに、右の表にあるようにChastityの値が0.6以下のサイクルが2つ以上あった場合、Filterを通さない。



$$6/(6+5) = 0.54$$

Chastityの計算式

Summary Tab の中のどこに%Clusters Passing Filterがあるのか

illumina Sequencing Analysis Viewer 1.8.11 - Christian Hans

Sequencing Analysis Viewer

Run Folder: C:\Documents and Settings\Desktop\%PF dataset Browse Refresh

Analysis | Imaging | **Summary** | Tile Status | TruSeq Controls | Indexing

Run Summary

Level	Yield Total (G)	Projected Total Yield (G)	Yield Perfect (G)	Yield <=3 errors (G)	Aligned (%)	% Perfect [Num Cycles]	% <=3 errors [Num Cycles]	Error Rate (%)	Intensity Cycle 1	% Intensity Cycle 20	% >= Q30
Read 1	0.0	0.0	0.0	0.0	0.00	0.0 [20]	0.0 [20]	0.00	396	111.8	0.6
Read 2 (I)	0.0	0.0	0.0	0.0	0.00	0.0 [5]	0.0 [5]	0.00	831	0.0	0.1
Read 3	0.1	0.1	0.0	0.0	0.01	0.0 [43]	0.0 [43]	0.00	155	239.3	56.8
Total	0.1	0.1	0.0	0.0	0.01	0.0	0.0	0.00	461	175.5	35.1

Read 1

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1	% Intensity Cycle 20
1	12	973 +/- 61	18.29 +/- 4.94	1.209 / 1.754	7.94	1.45	0.6	0.0	0	0.0 +/- 0.0	0.00 +/- 0.00	0.00 +/- 0.00	0.00 +/- 0.00	0.00 +/- 0.00	396 +/- 39	111.8 +/- 9.6

Read 2 (I)

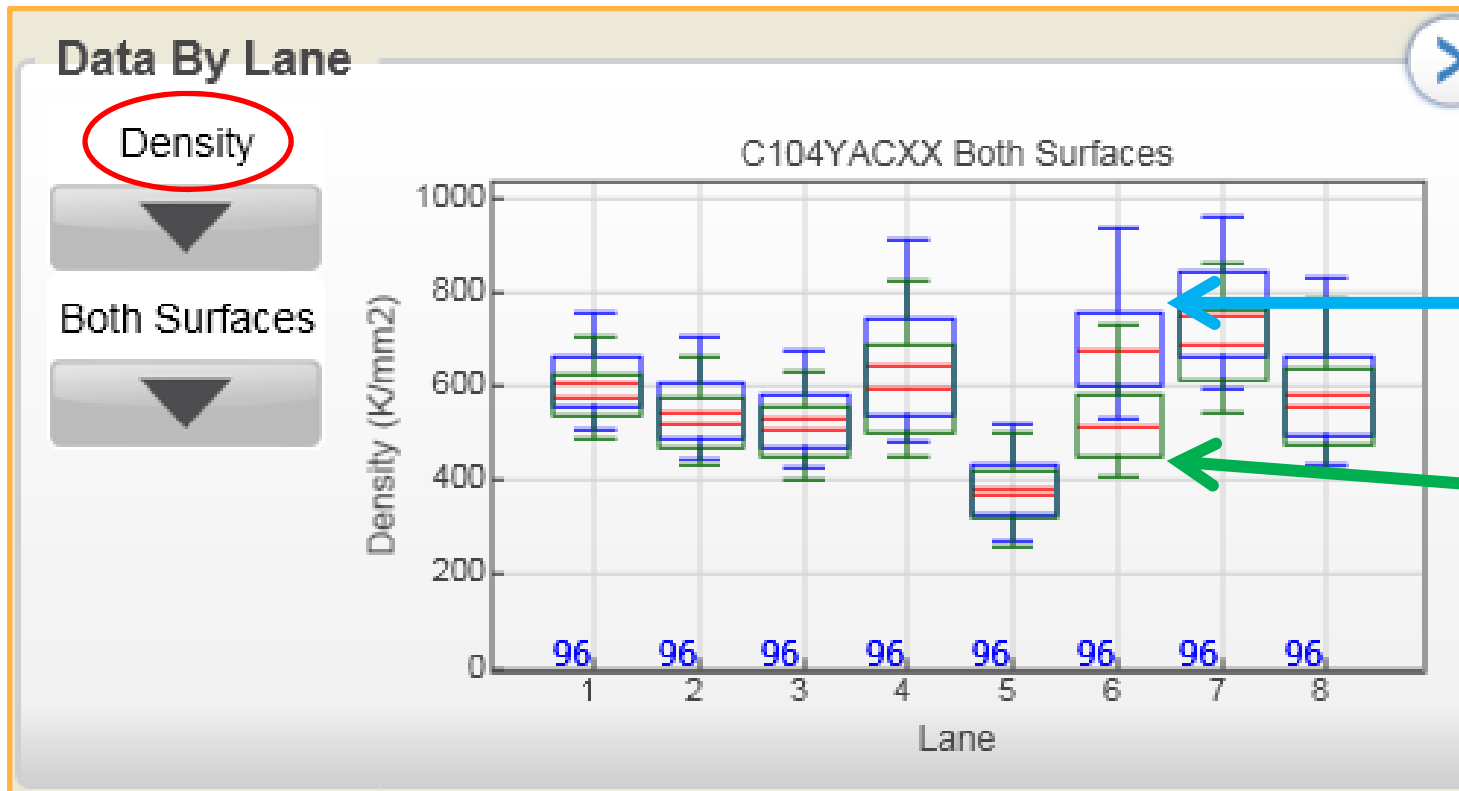
Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1	% Intensity Cycle 20
1	12	973 +/- 61	18.29 +/- 4.94	0.000 / 0.000	7.94	1.45	0.1	0.0	0	0.0 +/- 0.0	0.00 +/- 0.00	0.00 +/- 0.00	0.00 +/- 0.00	0.00 +/- 0.00	831 +/- 102	0.0 +/- 0.0

Read 3

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1	% Intensity Cycle 20
1	12	973 +/- 61	18.29 +/- 4.94	0.217 / 0.217	7.94	1.45	56.8	0.1	0	0.0 +/- 0.0	0.00 +/- 0.00	0.00 +/- 0.00	0.00 +/- 0.00	0.00 +/- 0.00	155 +/- 15	239.3 +/- 40.2

Copy to Clipboard... Generate IVC Plots...

Data By Laneのどこに Clusters Passing Filterがあるのか



Box Plots	
Blue	Raw Cluster Counts
Green	Clusters passing filter

GTATCATAGATATTTTATCCACTGATTCAGGTTACCGTATCCGAGGAGTATCAATTDAGACTAARATTAACGTACCATTAAGAGGTCACGTTGCGAAGCGCGGAAAGGAAATGATAAATTAACACACTTCTTTAAGCTTAAAGGGAAGGTATCATTAAGATTAATCT
ACTTACACACTTTGTTAGGCTTAAGATTCCTTGTCCACTGATTCAGGTTACCGTATCCGAGGAGTATCAATTDAGACTAARATTAACGTACCATTAAGAGGTCACGTTGCGAAGCGCGGAAAGGAAATGATAAATTAACACACTTCTTTAAGCTTAAAGGGAAGGTATCATTAAGATTAATCT
GCTTGCACACTGAGCTTGTGTTAGCTTAAATTAATTTGATCCACTGATTCAGGTTACCGTATCCGAGGAGTATCAATTDAGACTAARATTAACGTACCATTAAGAGGTCACGTTGCGAAGCGCGGAAAGGAAATGATAAATTAACACACTTCTTTAAGCTTAAAGGGAAGGTATCATTAAGATTAATCT
CTTTCCTTAAGATTCCTTAACTCCACTGATTCAGGTTACCGTATCCGAGGAGTATCAATTDAGACTAARATTAACGTACCATTAAGAGGTCACGTTGCGAAGCGCGGAAAGGAAATGATAAATTAACACACTTCTTTAAGCTTAAAGGGAAGGTATCATTAAGATTAATCT
CAATCCACTTAAAGATTCCTTAACTCCACTGATTCAGGTTACCGTATCCGAGGAGTATCAATTDAGACTAARATTAACGTACCATTAAGAGGTCACGTTGCGAAGCGCGGAAAGGAAATGATAAATTAACACACTTCTTTAAGCTTAAAGGGAAGGTATCATTAAGATTAATCT
TTATCCACTTAAAGATTCCTTAACTCCACTGATTCAGGTTACCGTATCCGAGGAGTATCAATTDAGACTAARATTAACGTACCATTAAGAGGTCACGTTGCGAAGCGCGGAAAGGAAATGATAAATTAACACACTTCTTTAAGCTTAAAGGGAAGGTATCATTAAGATTAATCT
GTATCATTAAGATTCCTTAACTCCACTGATTCAGGTTACCGTATCCGAGGAGTATCAATTDAGACTAARATTAACGTACCATTAAGAGGTCACGTTGCGAAGCGCGGAAAGGAAATGATAAATTAACACACTTCTTTAAGCTTAAAGGGAAGGTATCATTAAGATTAATCT



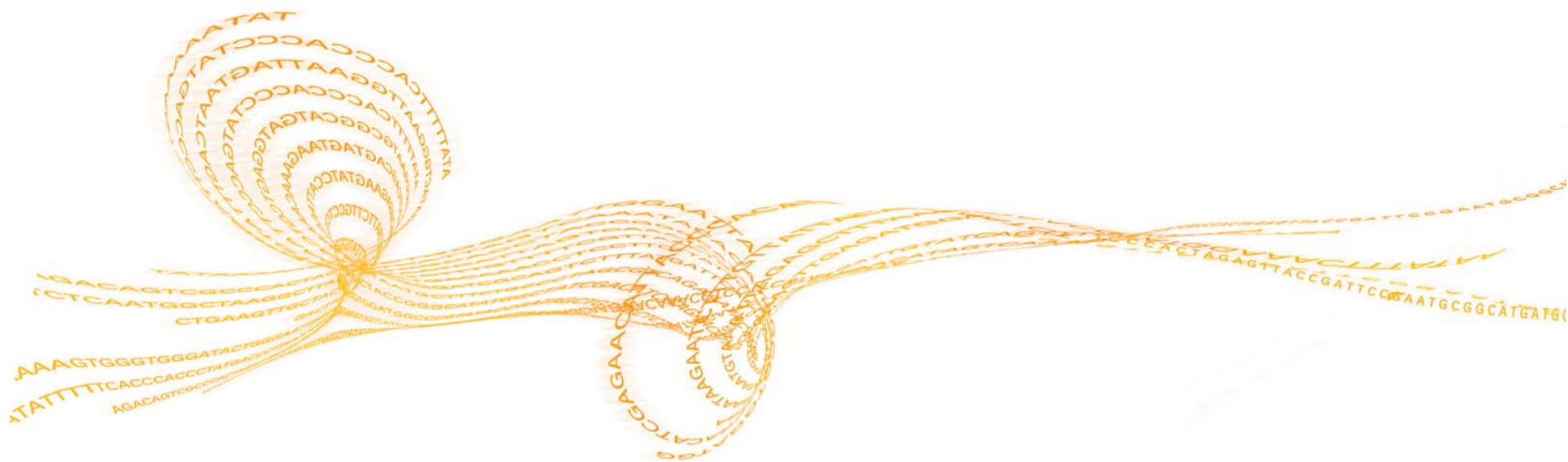
- ▶ %Clusters Passing Filter (%PF)の定義
- ▶ どのくらいの %PFが得られるのか
- ▶ %PFが低くなる主な原因とその診断方法
- ▶ %PFを最適化するには

どのくらいの %Clusters Passing Filter が得られるのか

PhiX コントロールDNAを用いて推奨されているCluster Density以内のクラスターを作成した場合

Read 1				
Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)
1	96	611 +/- 61	95.40 +/- 0.68	0.216 / 0.272
2	96	551 +/- 67	95.79 +/- 0.74	0.214 / 0.273
3	96	533 +/- 64	95.66 +/- 1.81	0.210 / 0.270
4	96	657 +/- 118	92.59 +/- 0.96	0.174 / 0.204
5	96	388 +/- 61	96.57 +/- 0.25	0.196 / 0.226
6	96	691 +/- 98	76.40 +/- 1.20	0.208 / 0.274
7	96	757 +/- 103	91.32 +/- 0.91	0.191 / 0.222
8	96	589 +/- 104	96.38 +/- 0.65	0.245 / 0.201

% PF は80以上



- ▶ %Clusters Passing Filter (%PF)の定義
- ▶ どのくらいの %PFが得られるのか
- ▶ %PFが低くなる主な原因とその診断方法
- ▶ %PFを最適化するには

%Clusters Passing Filter に影響を与える要因

- オーバークラスター
- 塩基の偏り
- Chemistry (装置もしくは試薬の問題)



HiSeq 2500
HiSeq 1500



HiSeq 2000
HiSeq 1000



MiSeq



Genome Analyzer IIX

GTATCATAGATATCTTATCCACTGATTCAGGTTACCGTAAACGGGANDGTATCAATDAGACTAAATAATAAGGTATCATTAAGAGGCTACCGTTCGAGCCGCGGAAABBAATGATAAAGTAAACACACTTCTTTACGCTTAAAGCGAAGGTATCATTAAGATTACT
ACTTACACAGCTTGTAGGCTTAAATATCTTGTCCACGATTCAGCTACCGTAAATATCTTGTATDCACTGATTCAGAGGTTACCGTTCAGGAACTTGTATGAGCTAAATATTAAGGTATCATTAAGAGGCTACCGTTCGAGCCGCGGAAABBAATGATAAAGTAAACACACTTCTTTACGCTTAAAGCGAAGGTATCATTAAGATTACT
GCTTGCACAGCTTGTAGGCTTAAATATCTTGTCCACGATTCAGCTACCGTAAATATCTTGTATDCACTGATTCAGAGGTTACCGTTCAGGAACTTGTATGAGCTAAATATTAAGGTATCATTAAGAGGCTACCGTTCGAGCCGCGGAAABBAATGATAAAGTAAACACACTTCTTTACGCTTAAAGCGAAGGTATCATTAAGATTACT
CTTCTTACGCTTAAAGCGAAGGTATCATTAAGATTACTTGTCCACGATTCAGCTACCGTAAATATCTTGTATDCACTGATTCAGAGGTTACCGTTCAGGAACTTGTATGAGCTAAATATTAAGGTATCATTAAGAGGCTACCGTTCGAGCCGCGGAAABBAATGATAAAGTAAACACACTTCTTTACGCTTAAAGCGAAGGTATCATTAAGATTACT
CAACGCTTAAAGCGAAGGTATCATTAAGATTACTTGTCCACGATTCAGCTACCGTAAATATCTTGTATDCACTGATTCAGAGGTTACCGTTCAGGAACTTGTATGAGCTAAATATTAAGGTATCATTAAGAGGCTACCGTTCGAGCCGCGGAAABBAATGATAAAGTAAACACACTTCTTTACGCTTAAAGCGAAGGTATCATTAAGATTACT
CTTCTTACGCTTAAAGCGAAGGTATCATTAAGATTACTTGTCCACGATTCAGCTACCGTAAATATCTTGTATDCACTGATTCAGAGGTTACCGTTCAGGAACTTGTATGAGCTAAATATTAAGGTATCATTAAGAGGCTACCGTTCGAGCCGCGGAAABBAATGATAAAGTAAACACACTTCTTTACGCTTAAAGCGAAGGTATCATTAAGATTACT
GTATCATAGATATCTTATCCACTGATTCAGGTTACCGTAAACGGGANDGTATCAATDAGACTAAATAATAAGGTATCATTAAGAGGCTACCGTTCGAGCCGCGGAAABBAATGATAAAGTAAACACACTTCTTTACGCTTAAAGCGAAGGTATCATTAAGATTACT

%Clusters Passing Filterが低くなる主な原因と その調査方法

原因: オーバークラスター
オーバークラスターの診断方法

1) Data by Lane

青と緑のボックスプロットの比較

オーバークラスターしているフローセルのCluster densityは
実際よりも低い数値となる

2) サムネイル画像の確認

推奨最大cluster density/mm²

- MiSeq

V2試薬キット: 880-965k/mm²

V3試薬キット: 1200-1400k/mm²

- HiSeq2000 : 750-850k/mm²

- GAIIx: 700-800k/mm²

%Clusters Passing Filterが低くなる主な原因と その調査方法

原因: オーバークラスター

推奨最大cluster density/mm²

- MiSeq
V2試薬キット: 880-965k/mm²
V3試薬キット: 1200-1400k/mm²

Box Plots	
Blue	Raw Cluster Counts
Green	Clusters Passing filter



クラスター数は正確に計算されているようだが、極端にPFが低い。
これはオーバークラスターを示唆している

%Clusters Passing Filterが低くなる主な原因と その調査方法

原因: オーバークラスター
オーバークラスターの診断方法

1) Data by Lane

青と緑のボックスプロットの比較

オーバークラスターしているフローセルのCluster densityは
実際よりも低い数値となる

2) サムネイル画像の確認

推奨最大cluster density/mm²

• MiSeq

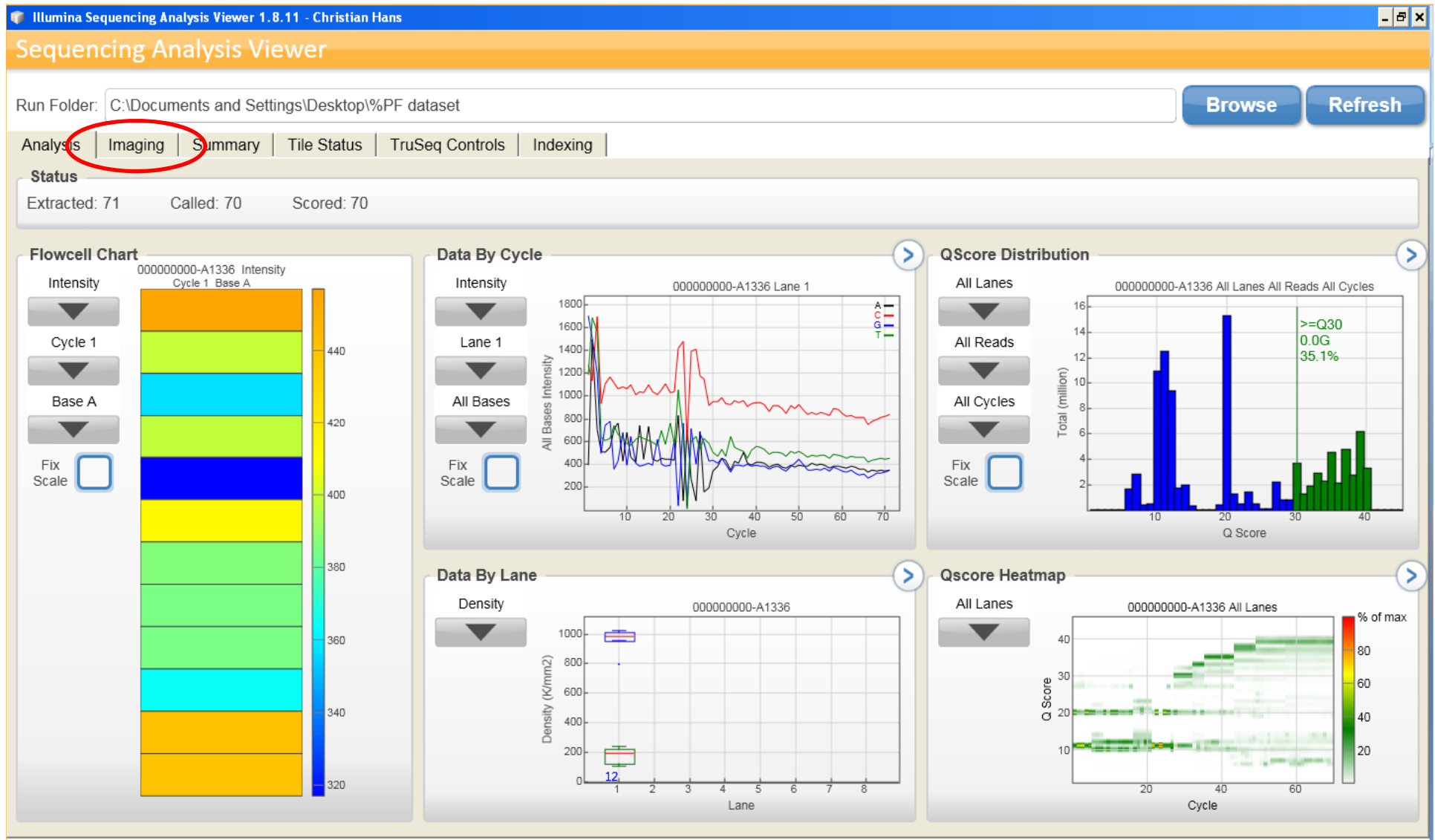
V2試薬キット: 880-965k/mm²

V3試薬キット: 1200-1400k/mm²

• HiSeq2000 : 750-850k/mm²

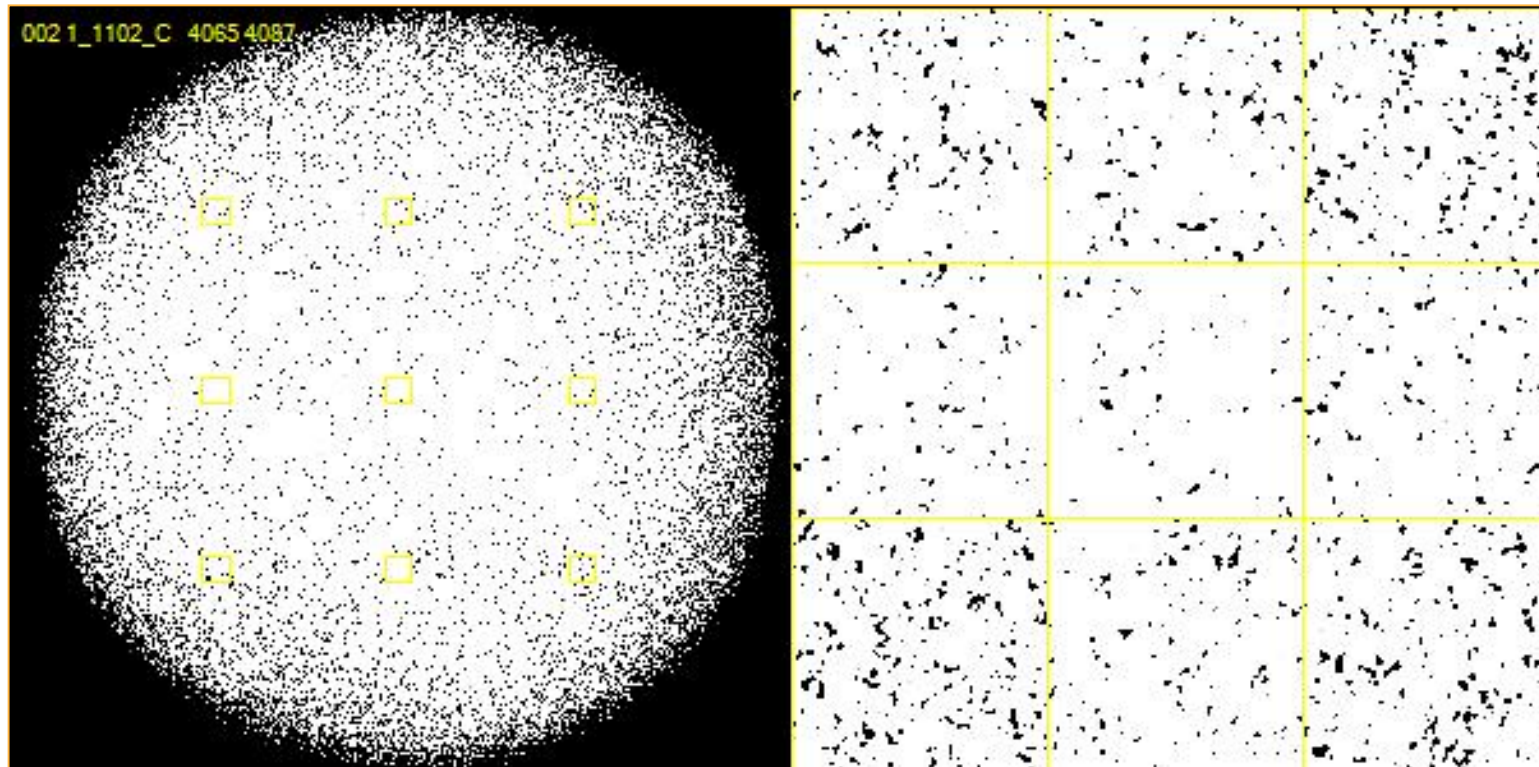
• GAIIx: 700-800k/mm²

Sequencing Analysis Viewer (SAV)



%Clusters Passing Filterが低くなる主な原因と その調査方法

原因: オーバークラスター



クラスターが飽和している状態

%Clusters Passing Filterが低くなる主な原因と その調査方法

原因: 塩基の偏り

塩基の偏りの診断方法:

- ▶ サンプルの偏りについての事前情報(カスタムバーコード、アンプリコン等)
- ▶ SAV: Data By Cycle (% base)

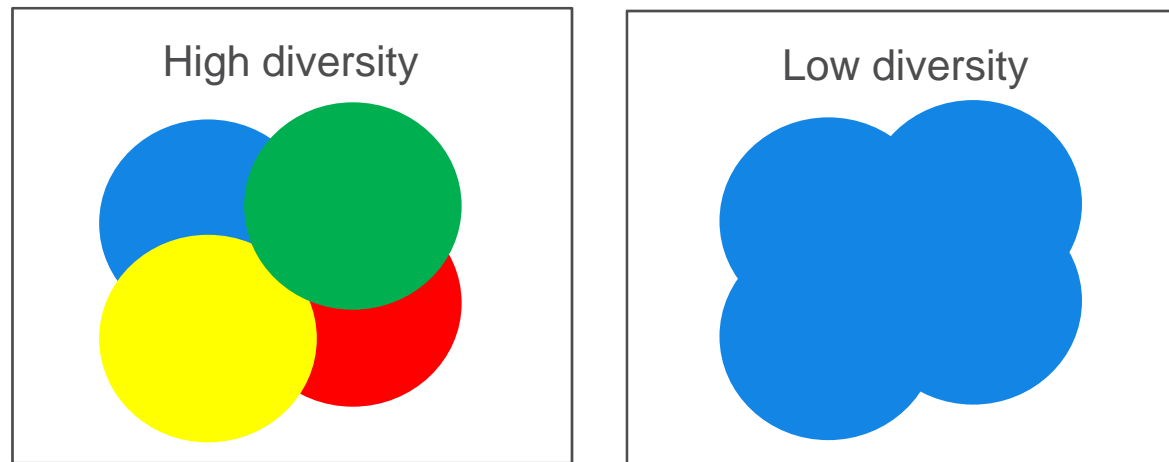
AAGCTCAGGCT AAGCTCAGGCT
AAGCTCAGGCT AAGCTCAGGCT AAGCTCAGGCT
AAGCTCAGGCT AAGCTCAGGCT
AAGCTCAGGCT AAGCTCAGGCT

%Clusters Passing Filterが低くなる主な原因と その調査方法

塩基に偏りがある

なぜ塩基の偏りがあると問題なのか:

- 近接したクラスターについて考えてみる



- 近接した別のクラスターから発せられるシグナルが同じ蛍光のために、近接したクラスターを別のクラスターとして区別することが困難になってしまう

%Clusters Passing Filterが低くなる主な原因と その調査方法

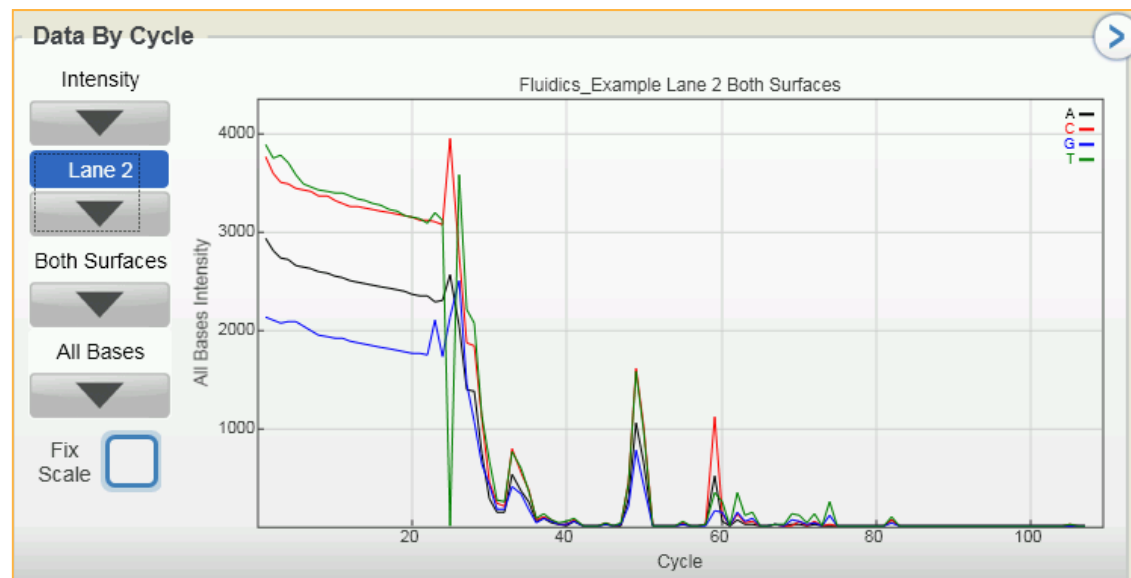
Chemistry (装置もしくは試薬の問題)

装置に問題がある場合の診断方法:

– Data By Cycle plot in SAV

Intensity plotで完全に数値が下がってしまう

流路の確認



%Clusters Passing Filterが低くなる主な原因とその調査方法

Chemistry (装置もしくはは試薬の問題)

症状:

- Q scoresが低くなる
- %PFが低くなる



試薬に問題がありそうなときは...

- 使用期限を確認する
- 調製方法を確認する
- 保存温度を確認する



- ▶ %Clusters Passing Filter (%PF)の定義
- ▶ どのくらいの %PFが得られるのか
- ▶ %PFが低くなる主な原因とその診断方法
- ▶ %PFを最適化するには

%Clusters Passing Filterを最適化するには

*Cluster density*を最適化する

定量が重要!!

-qPCR (*often the preferred method)

- 両端に適切なアダプターが付いたライブラリーのみが増幅する
- サンプルに似た性質を持つスタンダードを使用する



*ライブラリーサイズがブロードな場合はqPCRで正確に定量することができません。

ライブラリの定量方法

UV- 吸光度測定
(ex. Nanodrop)

- DNA、RNA、一本鎖、二本鎖の区別なく、核酸を非特異的に検出
- タンパク等がコンタミすると、測定値が大幅に上昇する (DNA量を過大評価する恐れがある)
- サンプル調製開始時、および完成したライブラリの定量には使わない。



Bioanalyzer
2100

- 定量の正確性はサンプルの希釈とハンドリングによって大きく左右する
- ライブラリの定性評価のみへの使用を推奨



二重鎖DNA特異的な蛍光定量法
(ex. Qubit or PicoGreen)

- 二重鎖DNAだけを特異的に検出
- 二重鎖DNAであれば、アダプター分子も検出する



定量PCR (qPCR)

- 完全な構造のライブラリのみを特異的に定量可能
- 検出・定量精度が非常に高い



%Clusters Passing Filterを最適化するには

塩基の偏りを減らす

ライブラリーへのPhiXのスパイク - インを行う

*注意: De novoの実験やPhiXゲノムと相関性が高いライブラリーにはご使用いただけません

PhiX はバクテリオファージ。

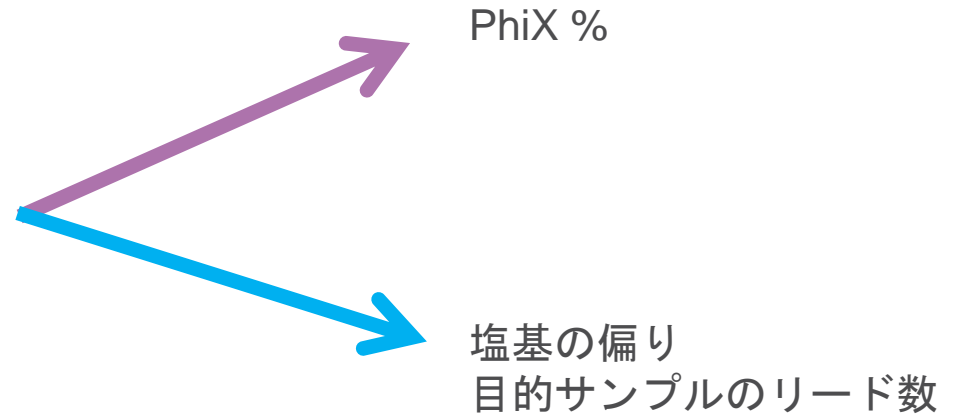
- 多様性がありランダムな配列 ⇒ 約 45% GC と 55% ATを含む
- スモールゲノム ⇒ アライメントやエラーレートの評価ができる
- 明確な定義 ⇒ リファレンスゲノムがしっかりと注釈付けされている

イルミナのPhiX コントロールライブラリーはすぐに使用できる

- アダプターが結合している
- 濃度は10nM (目的の濃度に希釈して使用)

%PFを最適化するには 塩基の偏りを減らす

どのくらいの PhiXが必要か？



- アウトプットとデータの質のバランスが重要
- 経験的に決める必要がある!

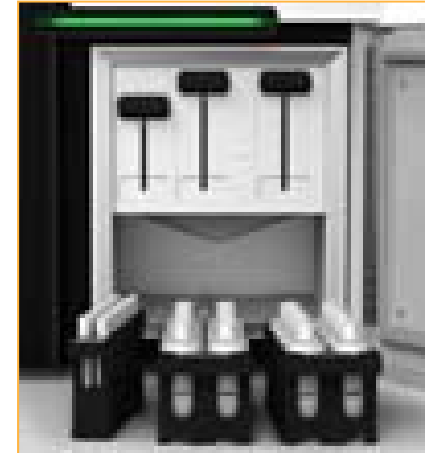
イルミナサポートウェビナー 2013/04/19
「塩基に偏りのある配列をMiSeqでシーケンスするには」
http://www.illumina.co.jp/events/webinar_japan.ilmn?ws=ss

イルミナサポートウェビナー 2013/03/22
「PhiXを使用してRun Qualityを改善する」
http://www.illumina.co.jp/events/webinar_japan.ilmn?ws=ss

%PFを最適化するには 流路の状態を最適にするには

- HiSeqのランを始める前に確認すること:
- メンテナンスウォッシュ時の廃液量
 - prime volume

- HiSeqのラン後に確認すること:
- 25サイクル目に算出される%PFとQ scores



試薬の適切な保存方法と手順を確認する

Summary

- ▶ %PFが低くなる主な原因
 - 1) オーバークラスター
 - 2) 塩基に偏りがある
 - 3) Chemistry (装置もしくは試薬)

- ▶ %PFの最適化に有効な手段
 - 1) *Cluster density*の最適化
 - 2) 塩基の偏りを減らす
 - 3) 流路の状態を最適にする

