

Feature Extraction

Version 9.5

User Training



内容

1. **Feature Extraction** のインストールとライセンス
2. **Feature Extraction** の機能: イメージ解析
 - **TIFF** イメージの解析
3. スポット数値化のためのセットアップとながれ
4. **Feature Extraction** の機能: スポット(数値化)
 - **Non-Agilent** アレイのグリッドあわせ
5. **Feature Extraction** スポット数値化アルゴリズム

 - スポット検出
 - フィーチャ&バックグラウンド領域の決定
 - シグナル強度の抽出
 - バックグラウンド補正
 - Multiplicative Detrending
 - 色素補正 (2 color)
 - 発現比計算 (2 color)
6. **QC**レポート
7. **Agilent miRNA**解析用マイクロアレイ(参考)
8. ソフトウェアのバージョン・アプリケーションごとの、**FE**プロトコルの対応表

Feature Extraction の インストールとライセンス

インストールについて

システム要件

PC Version

- Windows 2000 SP4 or later
Windows XP 32-bit
Windows Server 2003 SP1 *
Windows Vista 32-bit†
Windows Vista X64†‡
Windows XP X64†‡
- Pentium III 1.5 GHz or higher (Pentium IV 2.0 GHz or higher recommended)
- 1 GB RAM (2 GB recommended for high density (185K, 244K or higher))
- 20 GB available disk space
- 1024 x 768 display

* Feature Extraction software will run under Terminal Services but only 1 user will be allowed to run FE at any given time.

† Feature Extraction software installs under these operating systems, but currently QCChartTool and ExampleImages do not install.

‡ Feature Extraction software will not install under IA-64 Itanium Architectures, only 64-bit x86 (X64) architectures.

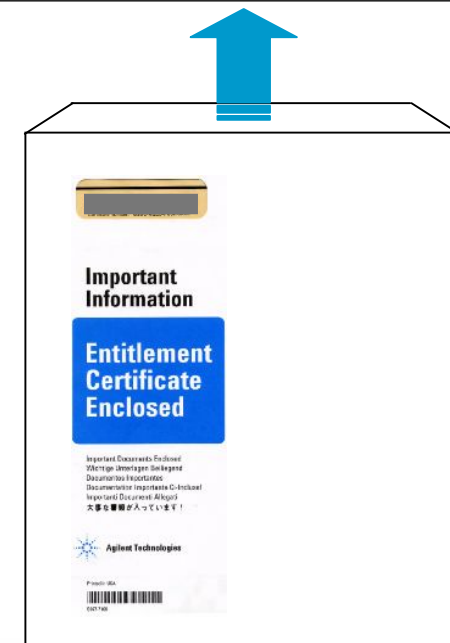
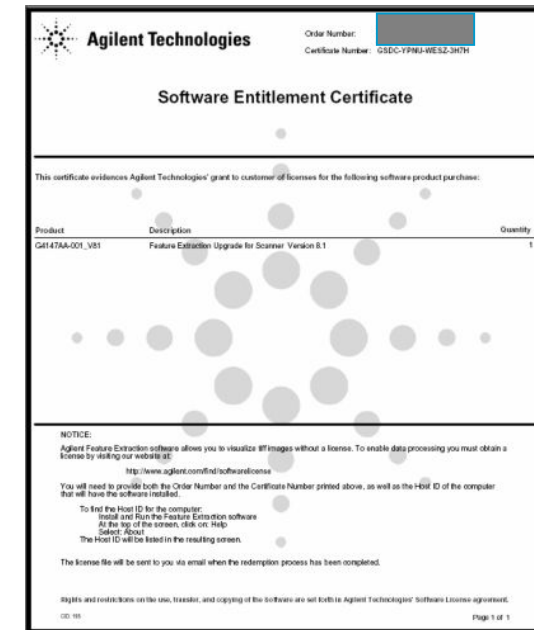
Administratorの権限でログインをする必要があります



ライセンスについて

2種類のライセンス

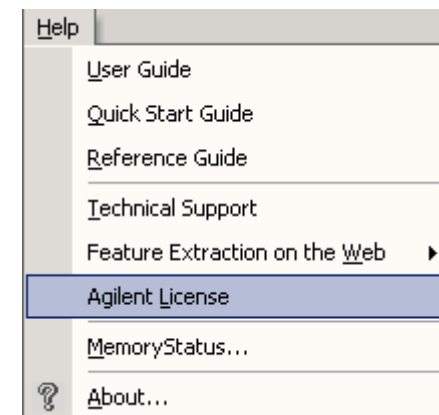
- 通常ライセンス
 - 1台の PC につき 1ライセンス使用
 - PC のホスト ID 情報が必要（PC は固定）
- 30日間デモ用ライセンス
 - Web からダウンロード可能
 - サポートはありません



ライセンスキーの取得法

- **Agilent** ウェブからライセンスキーを取得

<https://software.business.agilent.com>



あるいは Feature Extraction の Help > Agilent License を選択すると上記ウェブにリンクします

- ライセンスキーの取得には以下のものが必要

- **Order Number**

Feature Extractionソフトウェア2ライセンス権利証明書 (G2566-60000)に記載

- **Certificate Number**

Feature Extractionソフトウェア2ライセンス権利証明書 (G2566-60000)に記載

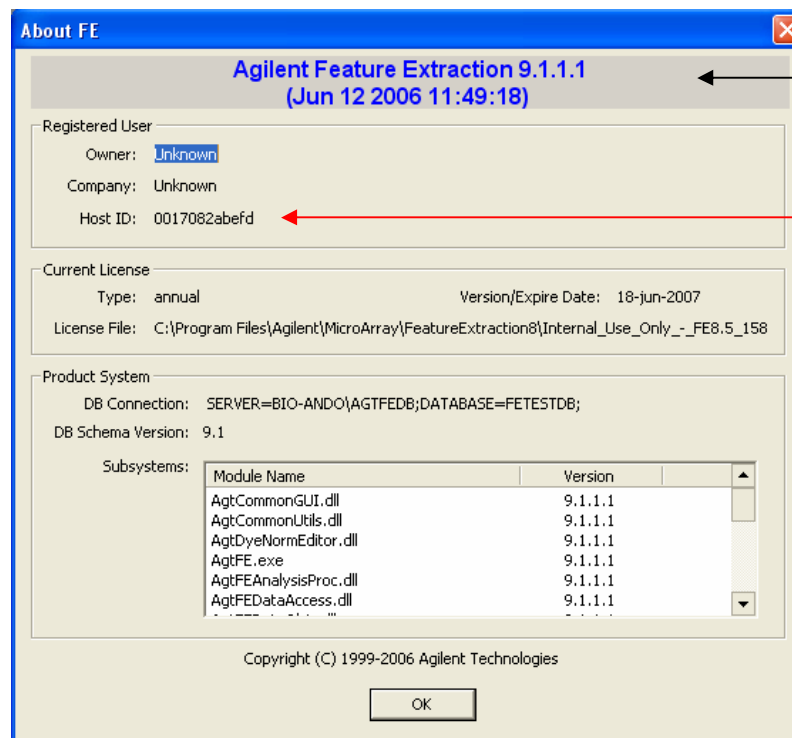
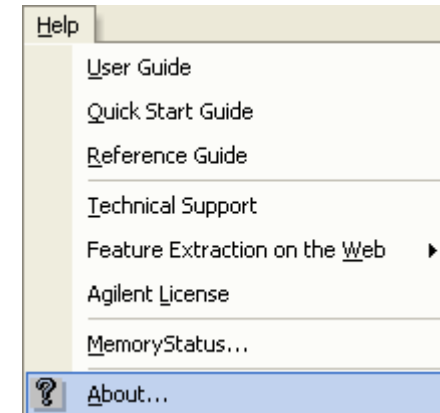
- **Host ID**

次スライド参照

- **メールアドレス**

ホストIDの調べ方

1. Feature Extraction をインストール
2. Help > About Analysis を選択
3. About Analysis ウィンドウ確認



ソフトウェアのバージョン

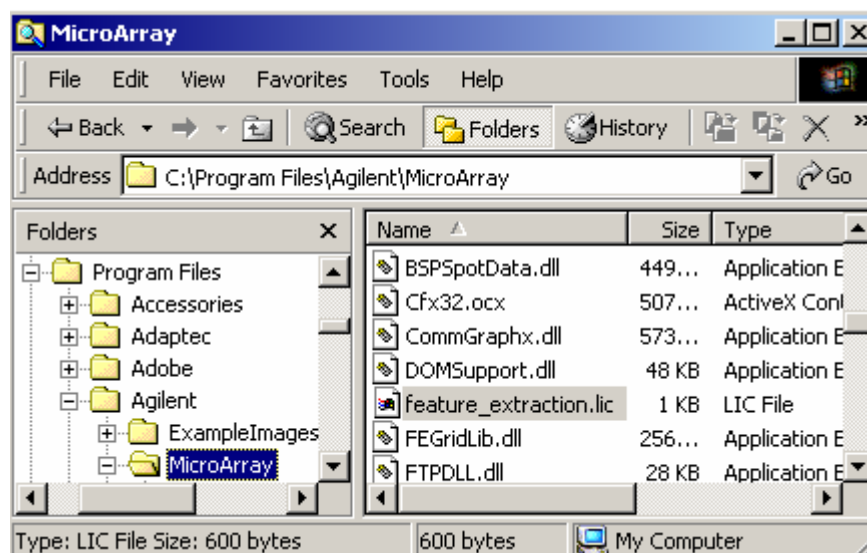
Host ID
(MAC/ Ethernet address)



ライセンスキーの保管

- ライセンスキー: **.lic**
- ライセンスキーは以下のディレクトリに保存

Program Files\Agilent\MicroArray



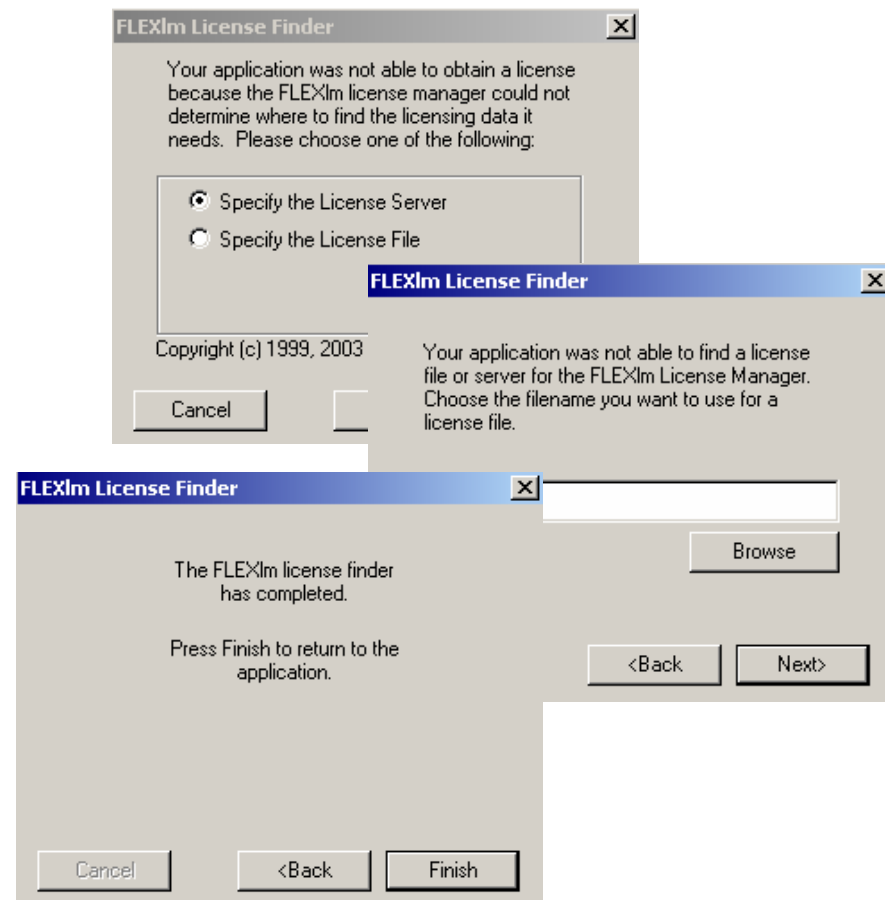
バックアップとしてライセンスキーは別の場所にも保管してください
(OS の再インストールの際にライセンスキーがなくなる可能性があります)

Feature Extraction ソフトウェアを使用する

インストール後、最初にライセンスキーを認識させる必要があります

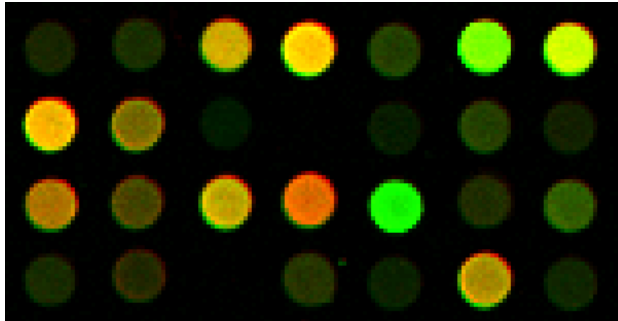
1. イメージファイルを開ける
2. **FLEXIm License Finder**
ウィンドウ
Specify the License File
を選択
3. 以下のディレクトリに保存した
ライセンスキーを選択

Program Files\Agilent\MicroArray



Feature Extraction の機能

Feature Extraction を使ったアレイ解析



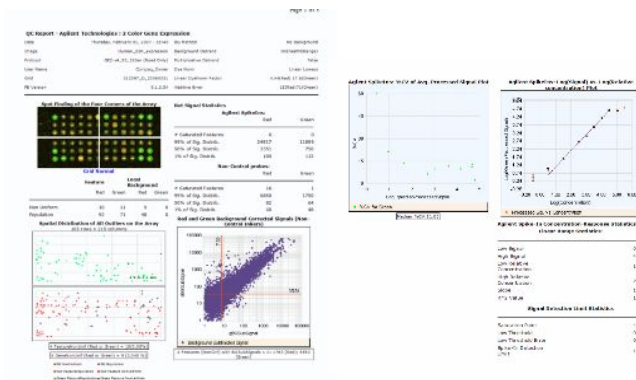
イメージ解析



スポットの数値化

+

QCレポート
(version 8より)



イメージ解析でできること

マイクロアレイのイメージを確認

- ・ イメージの色やスケールを変更
- ・ イメージのフリップ&回転(例:縦長→横長)
- ・ ヒストグラム、ラインプロットの作成
- ・ フィーチャとローカルバックグラウンドのフラグ確認
- ・ マイクロアレイ上のグリッド調整



スポット数値化でできること(1)

各スポット(フィーチャ)のシグナル抽出および発現比計算

- **Place Grid**
グリッドの初期化
- **Find and Measure Spots**
中心位置の決定→フィーチャとローカルバックグラウンド領域の決定→
ピクセルレベルのフラグ排除→シグナル強度計算
- **Flag Outliers**
フィーチャとローカルバックグラウンドのフラグづけ
- **Correct Background and Signal Biases**
バックグラウンド補正とシグナルの有意性検定
- **Correct Dye Biases**
色素バイアス補正のためのフィーチャ選びと補正
- **Compute Ratios and Errors**
ログ比(**Cyanine 5 / Cyanine 3**) およびエラー値、P値の計算



スポット数値化でできること(2)

スポット解析後のQCレポート

- **QC Report Header**
採用したバックグラウンド／色素補正法、バックグラウンドノイズ、色素補正係数
- **Spot Finding of Four Corners**
アレイの四隅のグリッド合わせと決定した中心位置を可視表示
- **Number and Spatial Distribution of Outliers and Up- / Down- Regulated Features**
フラグおよび発現差のあるフィーチャの数、およびこれらのフィーチャのアレイ上での分布
- **Net Signal Statistics and Local Background Inliers**
フィーチャおよびローカルバックグラウンドのシグナル値の統計
- **Foregournd Surface Fit**
フォアグラウンドの見積もり結果(アレイ表面のフィッティング)
- **Plot** : 赤と緑のバックグラウンド補正シグナル、**M-A**プロット

LogRatioの期待値と実測値.

バックグラウンド補正シグナルの**%CV**(スパイクイン)



Feature Extraction各バージョンでの変更点

FE8.5

- 1.グリッディング精度向上
- 2.Agilent遺伝子発現1色法データの数値化可能
- 3.アプリケーションごとのQCLレポート作成
- 4.Agilent2色法データのAxon画像数値化可能
- 5.新しいAdditive Detrendingの導入
Option: ネガコンRangeのFeatureを使う
- 6.新しいデザインファイルをeArrayから入手可能
- 7.FEのcommand line versionをサポート

FE9.1.3

- 1.Agilent高密度マイクロアレイの数値化可能
- 2.Agilentマイクロアレイのグリッディング精度向上
- 3.AgilentスキャナのXDRスキャンで得た画像ペアから数値化可能
- 4.Agilentマイクロアレイの数値化速度向上
- 5.1つのFeature Extraction protocolで、全てのフォーマットの数値化可能
- 6.Compactのテキストファイル出力可能
- 7.QC chart toolの導入と、QC metricsのFE内での使用可能

FE9.5.1

- 1.miRNAマイクロアレイの数値化が可能
- 2.QCLレポートがPDFファイル形式に変更
- 3.マルチプレックスアレイのグリッディング精度向上
- 4.全てのAgilentに対してグリッドテスト改善
- 5.Dye Norm files扱い改善とdye norm Editor改善
- 6.Agilent optimized QC metric setsの導入

FE9.5.3

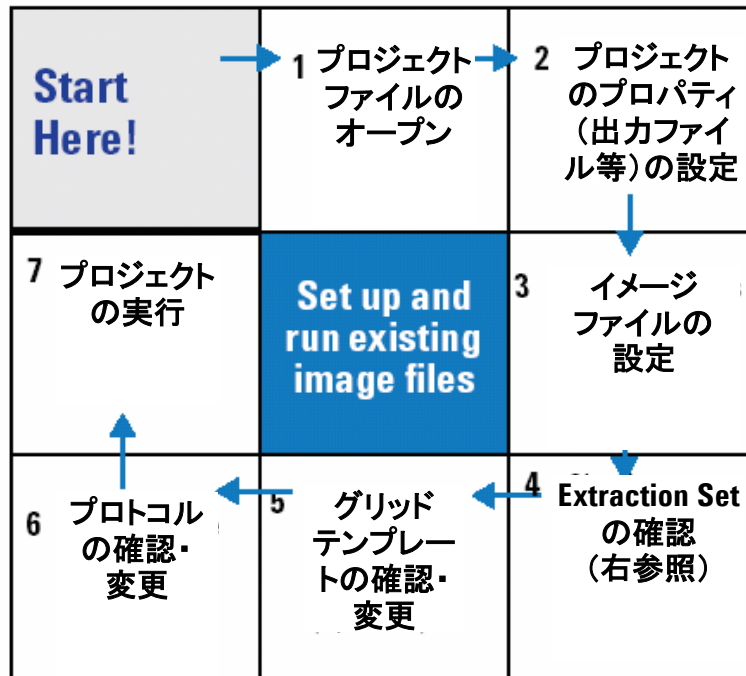
- 1.データベースエンジンを MSDEから
SQL Expressへ変更
- 2.Windows installerのアップデートにより、
Windows Server 2003, Windows XP 64-bit,
Windows Vista、Windows Vista 64-bitで動作
- 3.Windows Terminal ServicesでのFEの使用を
制限つきでサポート
- 4.QCLレポートの形式を、HTMLとPDFから選択可能
- 5.FE 9.5.3 の出力数値はFE 9.5.1と同一



FEプロジェクトの概念

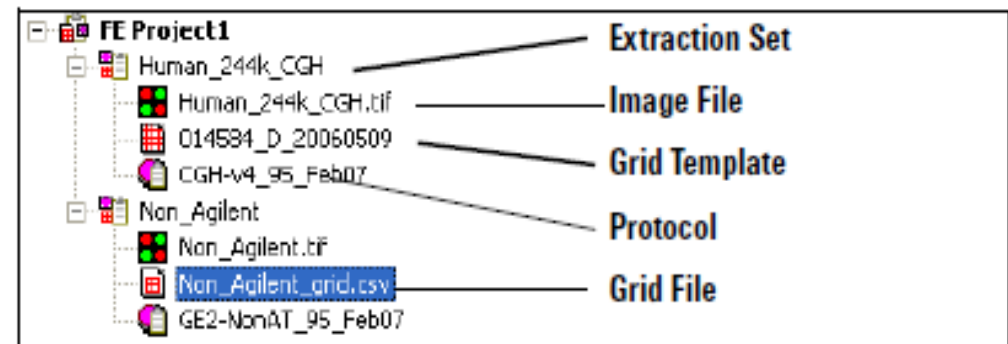
スポット数値化のバッチ処理

プロジェクトの作成から実行まで



(Standard Project)

FE Project と Extraction Set



Standard プロジェクト vs. On-Time プロジェクト

Standardプロジェクト

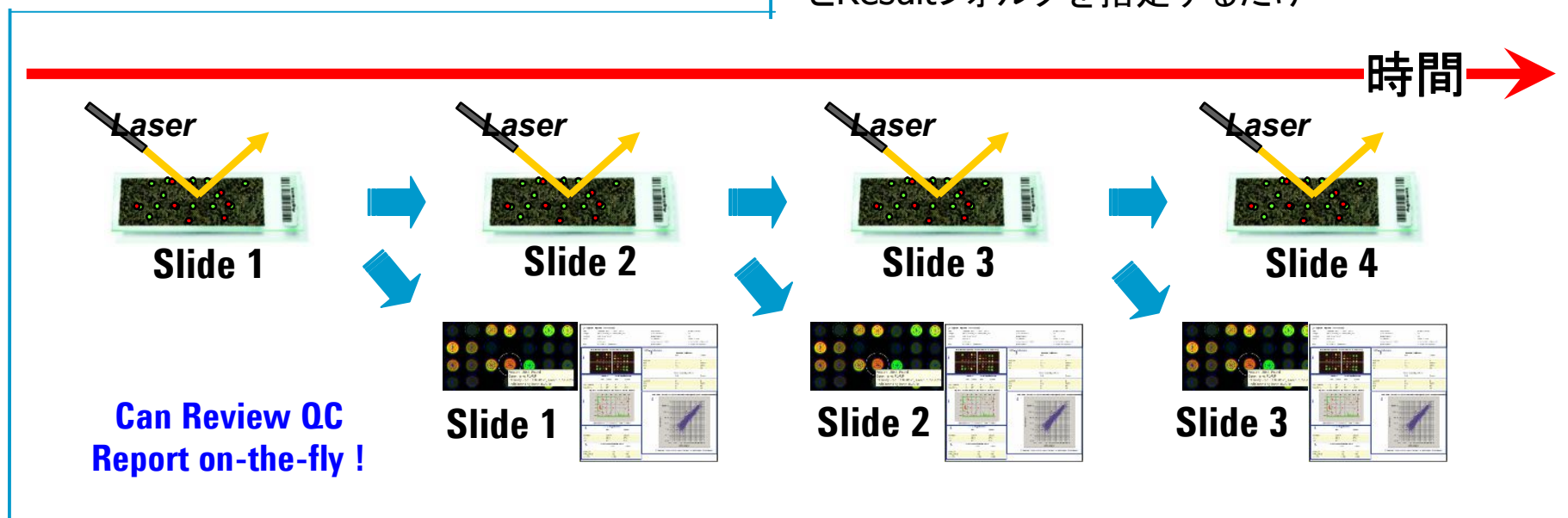
スキャンの終了した複数のTIFFイメージ(様々なフォーマット可)を一つのプロジェクト内で解析

On-Timeプロジェクト

スキャンをしながらスポット解析
("On-the-fly") – スキャンからスポット
解析まで完全自動化

→完全自動化にはバーコード情報が必要

→スキャン開始時にIncoming Imageフォルダ
とResultフォルダを指定するだけ



イメージ解析画面 ↔ スポット数値化画面 (FEプロジェクト)



イメージ画面への
切り替え(3通り)

1. ツールバーのイメージファイルオープンボタンをクリックまたは
File > Open > Image

2. TIFFイメージをワークエリアにドラッグ

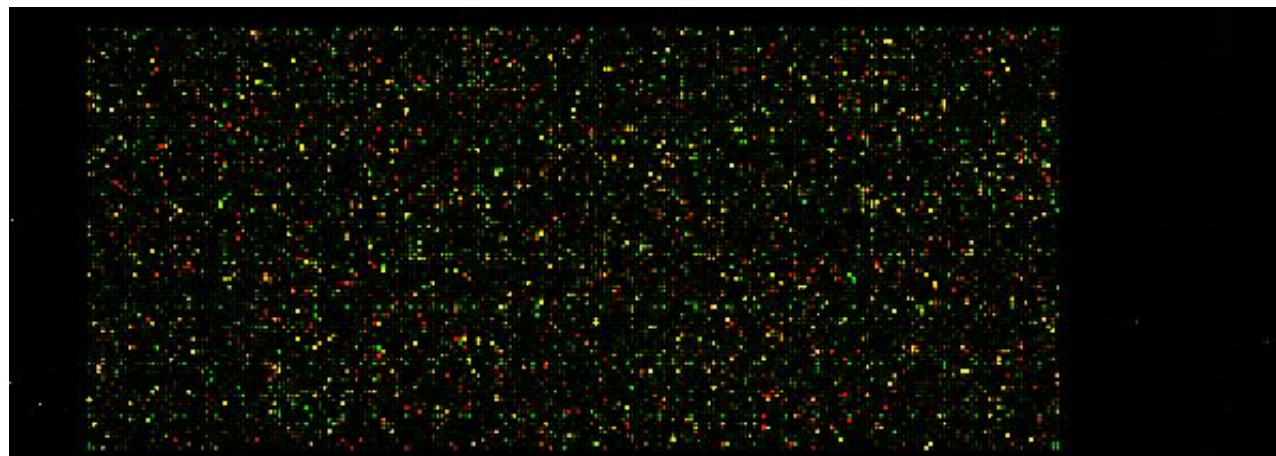
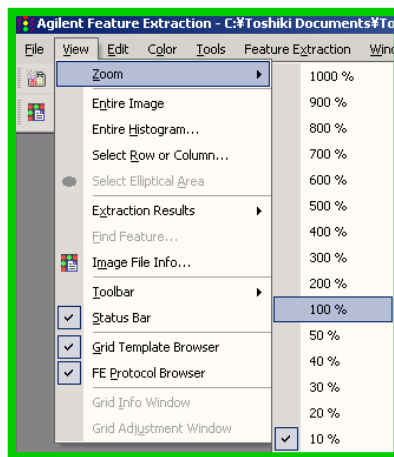
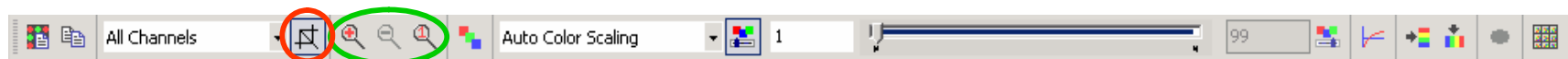
3. Project Explorerの
Extraction SetのTIFFイメージ
アイコンをダブルクリック



Feature Extraction

イメージ解析

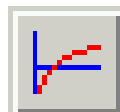
クロップモード、Zoom In & Out



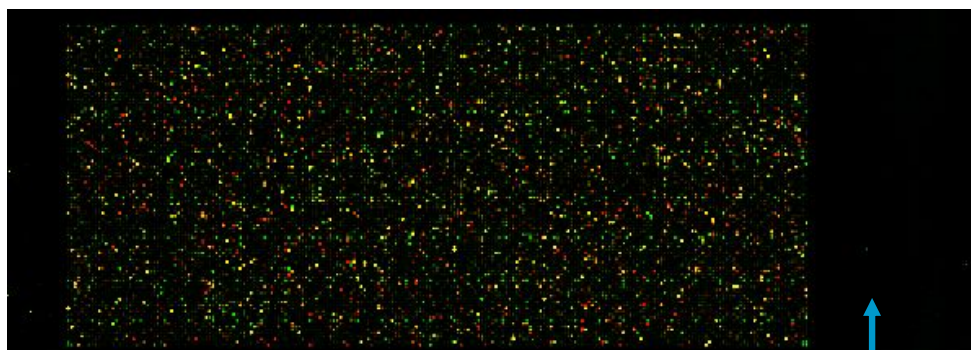
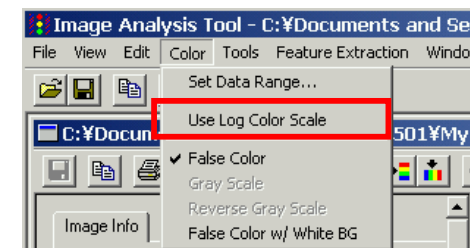
クロップモード**ON** →イメージをクロップ時に新しいウィンドウで表示

クロップモード**OFF** →イメージをクロップ時にその大きさにあわせてウィンドウを拡大

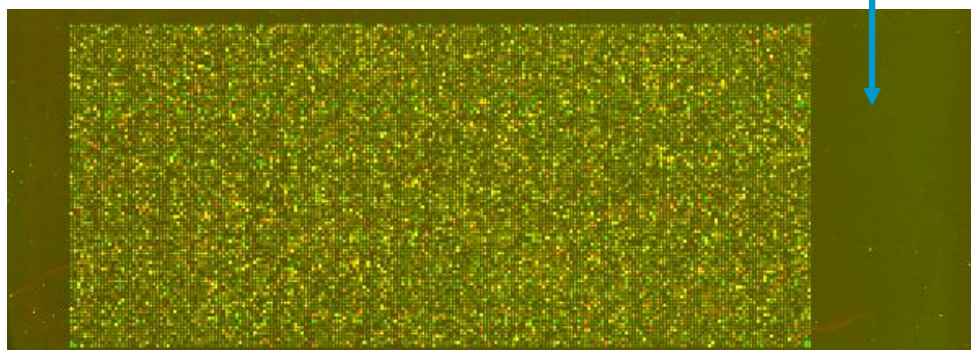
-  拡大  縮小  **100%** (PCによって異なります)
- **View > Zoom**
- ショートカット **Ctrl + (左ダブルクリック)** 拡大
 Ctrl + (右ダブルクリック) 縮小



スケールの変更（リニア/ログ）



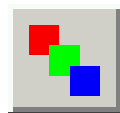
リニア表示



ログ表示

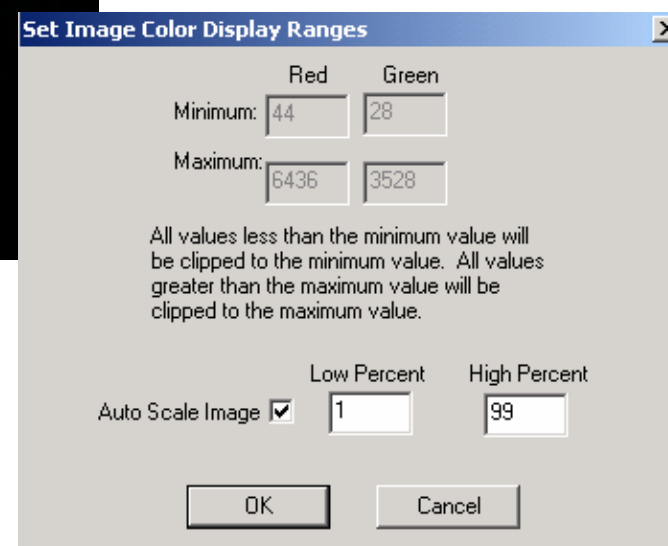
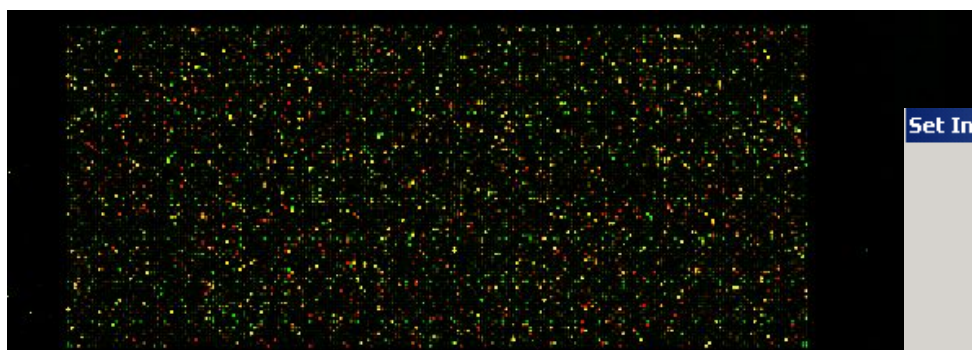
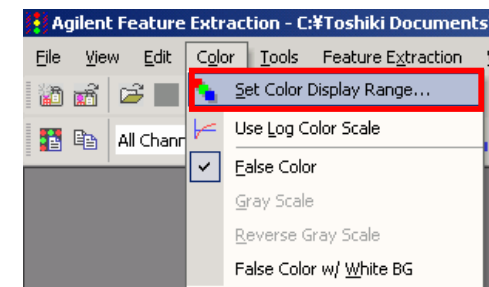
ログ表示にすることにより、アレイの大きさ、バックグラウンドのむらなどを確認





シグナル範囲

両チャンネルのシグナル範囲を確認

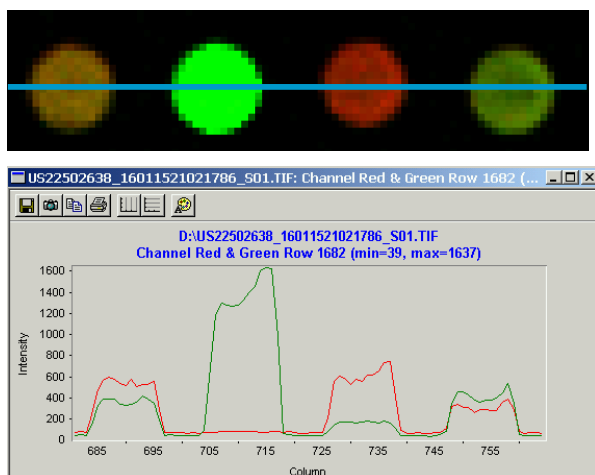
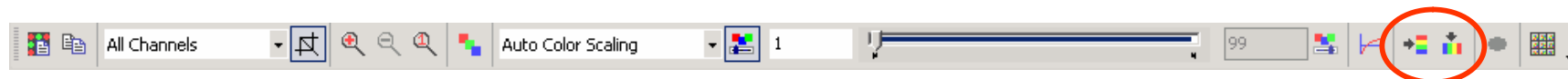


Set Image Color Display Ranges 内の数値またはスケーリングバーを変更することで、イメージ表示を変更可能



ラインプロット

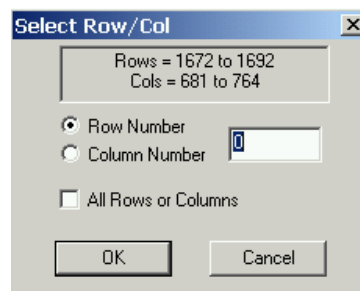
指定したラインでのシグナル値を確認



水平線



垂直線



あるいは左ダブルクリック

あるいは右ダブルクリック

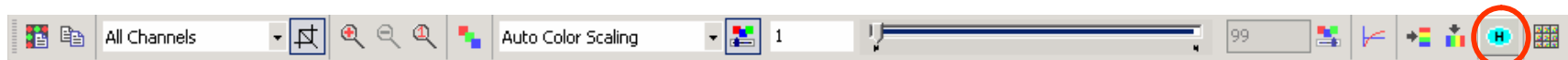
注意:ダブルクリックでラインプロットを確認する場合、
クロップモードはOFFにします





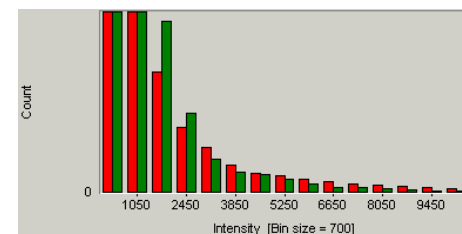
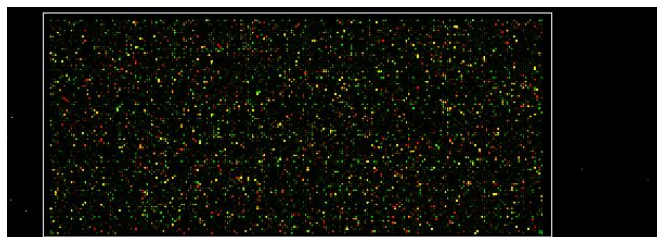
ヒストグラム

選択範囲のピクセルを使い、ヒストグラムを作成



クロップモードを**OFF**にする

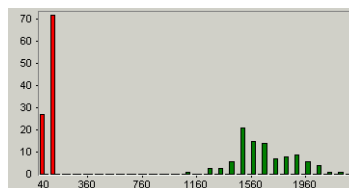
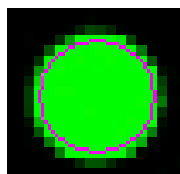
- 選択範囲＝四角（アレイ全体など）
右クリックドラッグでアレイの範囲を選択



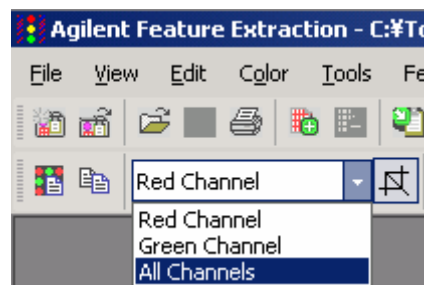
- 選択範囲＝丸型（個々のスポットなど）



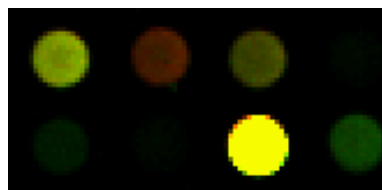
をクリックし、右クリックドラッグでスポットを範囲を選択



表示カラーの変更



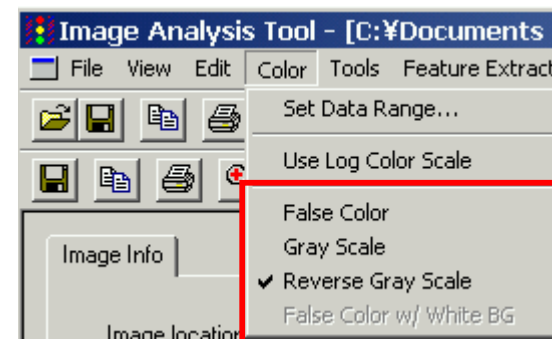
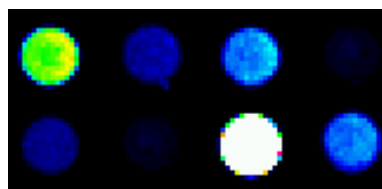
2色



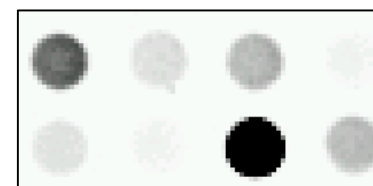
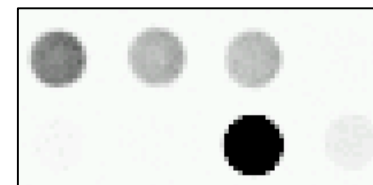
1色
(赤)



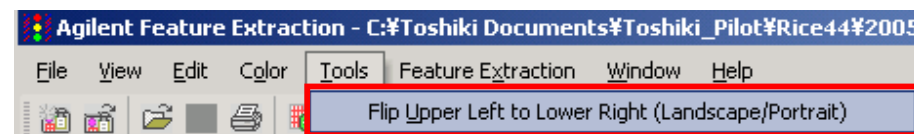
1色
(緑)



その他カラー表示の変更

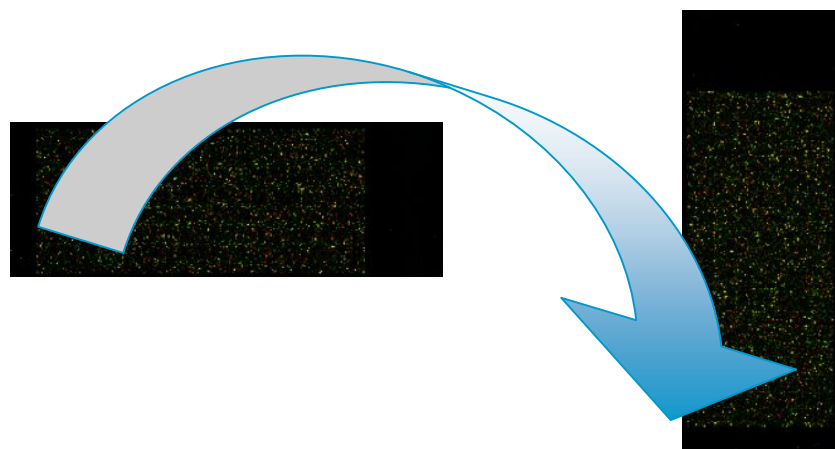


Tools メニュー



Flip Upper Left to Lower Right

アレイイメージを裏返し、縦横変換



例: Agilentアレイ
DNA裏面からよみとり
横長、バーコード左

例: Agilentアレイ
DNA表面からよみとり
縦長、バーコード下

Preferences...

イメージ解析、グリッド設定、FEプロジェクトなど、様々なユーザー設定が可能

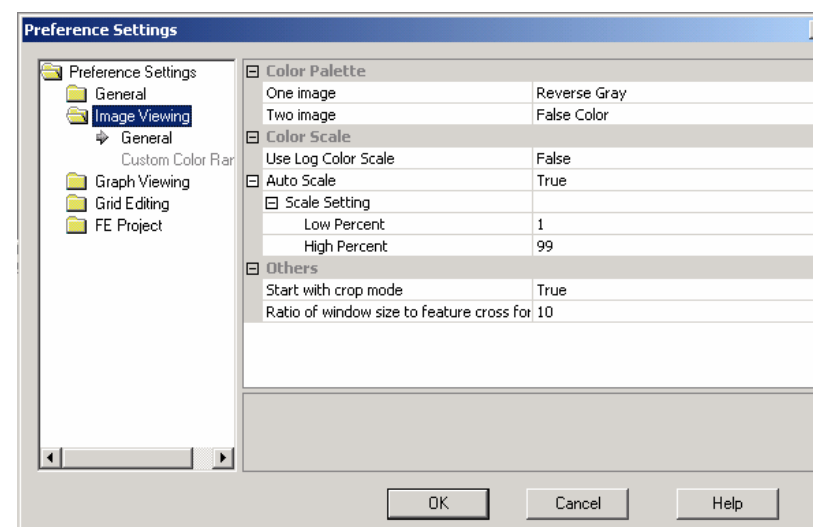
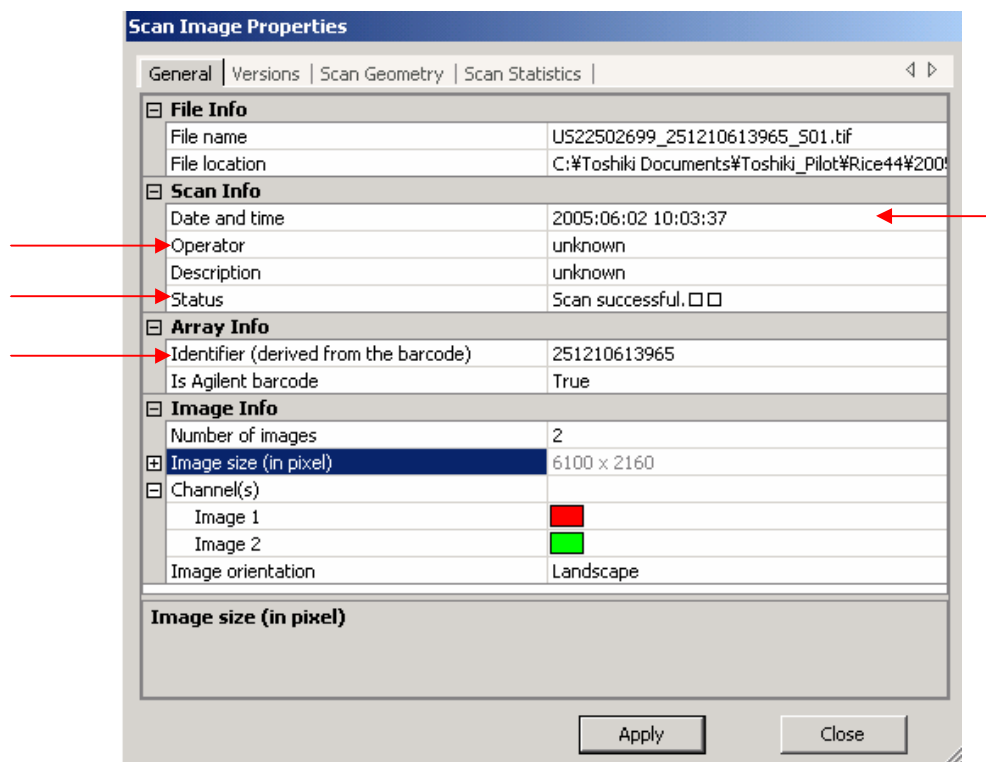
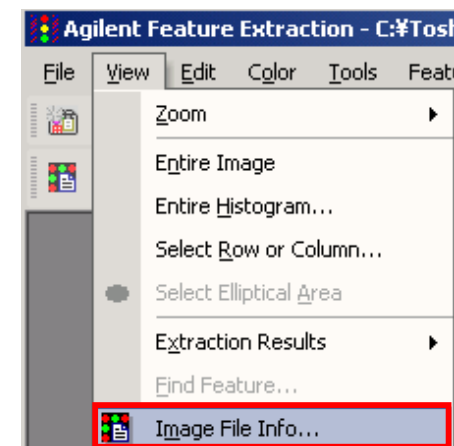




Image File Info

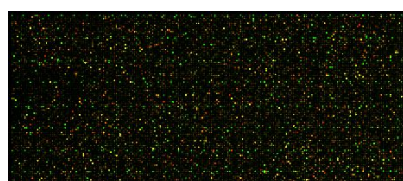
- スキャン時の情報
(Date&Time、Operator、スキャナの状態)
- その他バーコード情報、イメージ情報を確認



グリッドモードによる **Non-Agilent**マイクロアレイの数値化

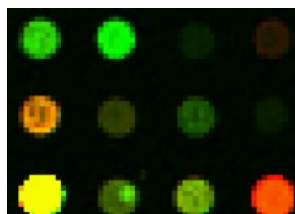
Feature Extraction 数値化までのステップ

Agilent アレイ



**Feature Extraction
Project 実行**

Non-Agilent アレイ



グリッド作業



**Feature Extraction
Project 実行**

Non-Agilent アレイグリッド作業: 言葉の定義

Input ファイル

- **Gene List** アレイの行列情報及びアノテーション情報が記載されているリスト
(ユーザー作成、数値化前に既にわかっている情報)

作業

- **グリッド** アレイ上のスポット中心位置の決定

Output ファイル

- **グリッドファイル** メイングリッド、サブグリッドの位置情報
- **フィーチャファイル** 各フィーチャの位置情報とアノテーション情報
(グリッドファイル作成時に自動的に作成される)

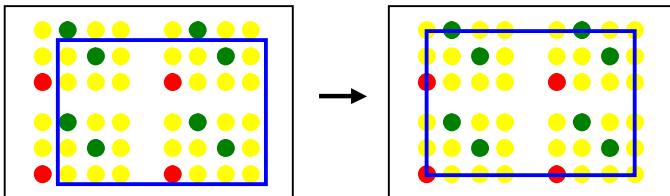
Non Agilent アレイの数値化ながれ

1. グリッドファイルの作成 使用するGene Listの選択

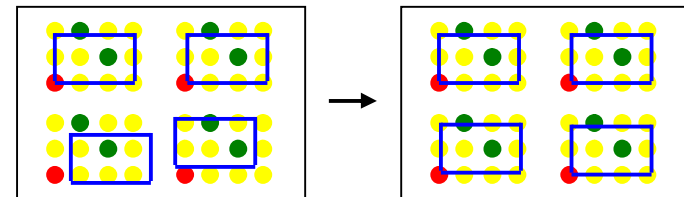
Gene list type:

- ☒ Agilent Grid File (*_grid.csv)
- ☐ GAL Format File (*.gal)
- ☐ Agilent Design File (*.xml)
- ☐ Tab Text File (*.t)
- ☐ No Gene List

2. メイングリッド位置の調整



3. サブグリッド位置の調整



4. スポット位置の調整



5. グリッドファイルの保存

6. Feature Extractor の実行

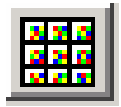
グリッドモード解析

- **Agilent** スキャナでイメージを読み取った **non-Agilent** マイクロアレイを数値化
- スポット位置検出ツールによって、アレイ上のスポット検出および位置を自動検出
- **Gene List** の形式 (アレイレイアウト、アノテーション情報)
 - **Agilent** グリッドファイル **.csv file**
 - **GAL** ファイル (**GenePix Array Layout**) **.gal file**
 - **Agilent** デザインファイル **.xml file**
 - タブ区切りテキストファイル **.txt file**

＜注意＞ **Agilent** スキャナはスライドガラスの裏側からマイクロアレイを読み取ります
以下の向きが一致していることをお確かめ下さい

- 遺伝子リストの並び順
- イメージの読み取り向き





グリッドの作成

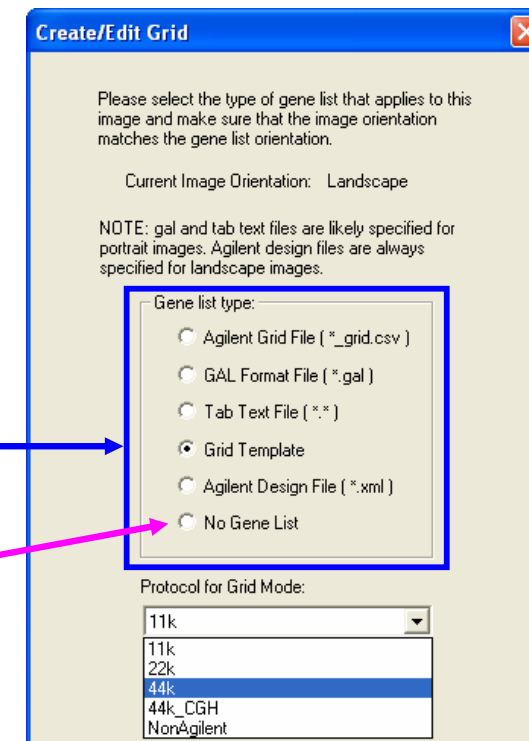
non-Agilentアレイを数値化するにはグリッドおよびフィーチャファイル作成が必要

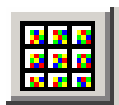
1. Grid Mode On/Off をクリックしグリッドモードにする



2. Create/Edit Grid 画面で使用する Gene List のタイプを選択

- **Gene List がある場合:**
_grid.csv, gal, xml, tab text
のいずれかを選択
- **Gene List がない場合:**
No Gene List を選択





グリッドの作成

3. 行列情報をもとに、各項目が **Geometry Information** ウィンドウに自動入力される

- **Gene List** がある場合はそのファイルの情報をもとに
- **Gene List** がない場合はアレイイメージから推測
- →どちらの場合もこの画面で数値の変更は可能

4. **Auto Fit** または **Manual Fit** をクリック

Geometry Information

Main grid / Array geometry

Dimensions

Columns: 2

Rows: 7

Nominal spacing between subgrids

Column (in pixels): 27.39

Row (in pixels): 30.08

Total number of subgrids: 14 (estimated)

Subgrid geometry

Dimensions

Columns: 30

Rows: 9

Nominal spacing between spots

Column (in pixels): 14.521

Row (in pixels): 15.025

Nominal spot diameter (in pixels)

Horizontal: 8.99

Vertical: 9.19

Maximum fit movements

☐ No limit for subgrid adjustment

☐ No limit for spot adjustment

Subgrid (in pixels)

Horizontal: 20

Vertical: 22

Spot location (in pixels)

Horizontal: 9

Vertical: 10

Reset

Auto Fit

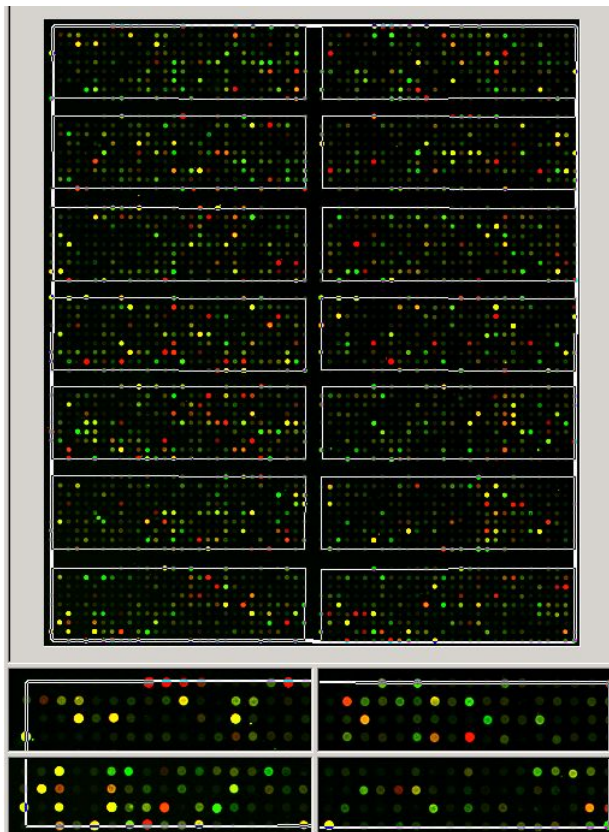
Manual Fit

Cancel

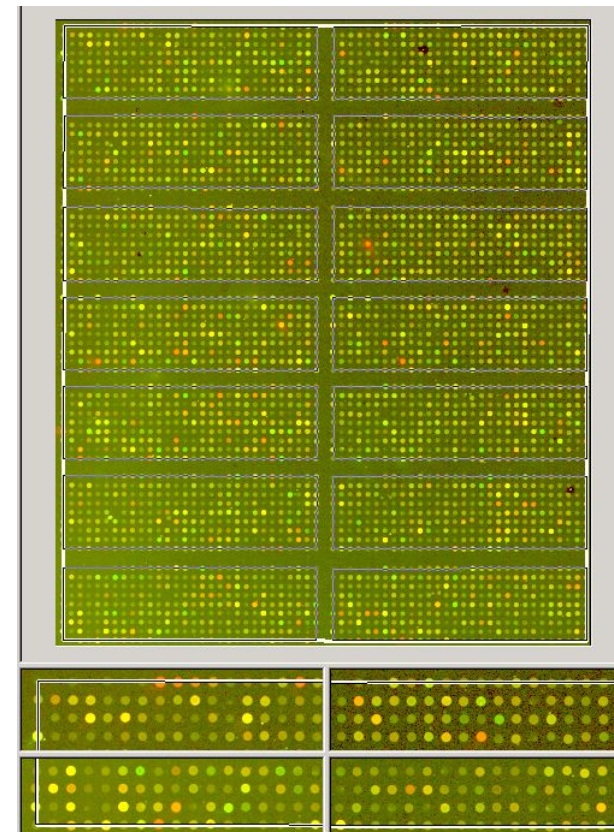


グリッドの調整

Color > Use Log Color Scale あるいは  を選択して
イメージをログ表示し、スポットを確認しやすいよう変換



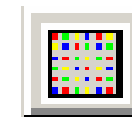
リニア表示



ログ表示

メイングリッドの調整

Adjust Main Grid



Adjust Main Grid をクリック

Adjust Main Grid

-  グリッドの位置調整

➤ Main Grid View

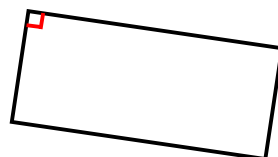
➤ Main Grid Corner View

Grid カーソルを使ってグリッドを移動し位置を調整

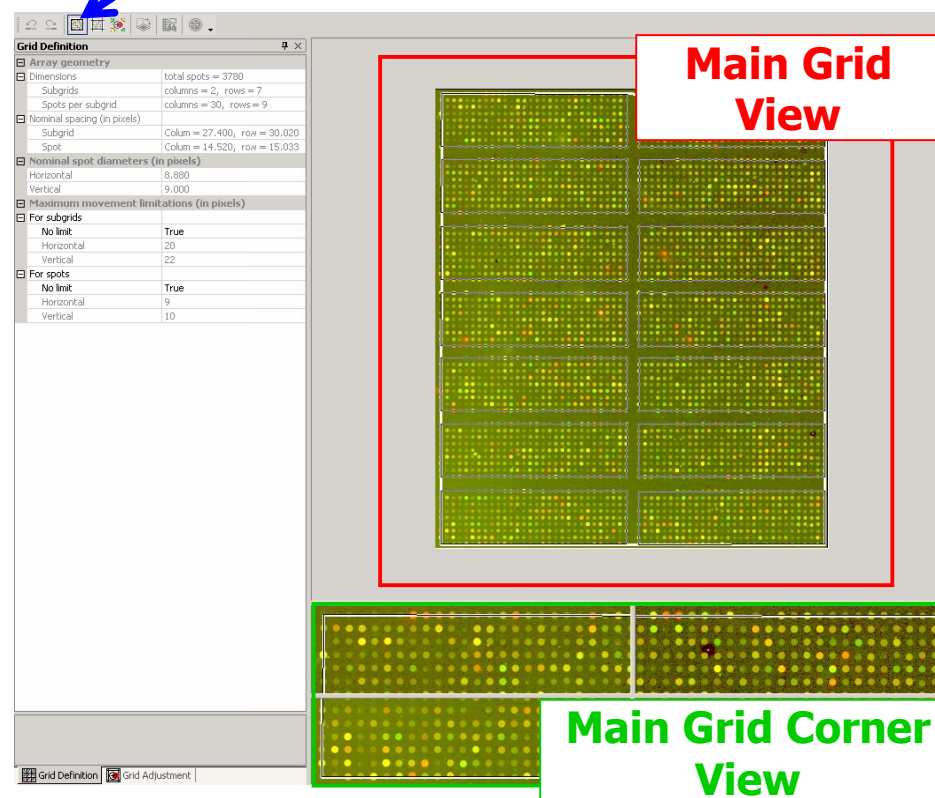
-  グリッドの回転

➤ Main Grid Corner View

マウスをアレイの隅にもっていき、
Rotate カーソルでメイングリッドを
回転

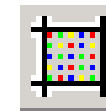


 Rotate



サブグリッドの調整

Adjust Sub Grid



Adjust SubGrid をクリック Adjust Subgrid Skew Subgrid

- サブグリッドの位置調整

➤ **Subgrid View**

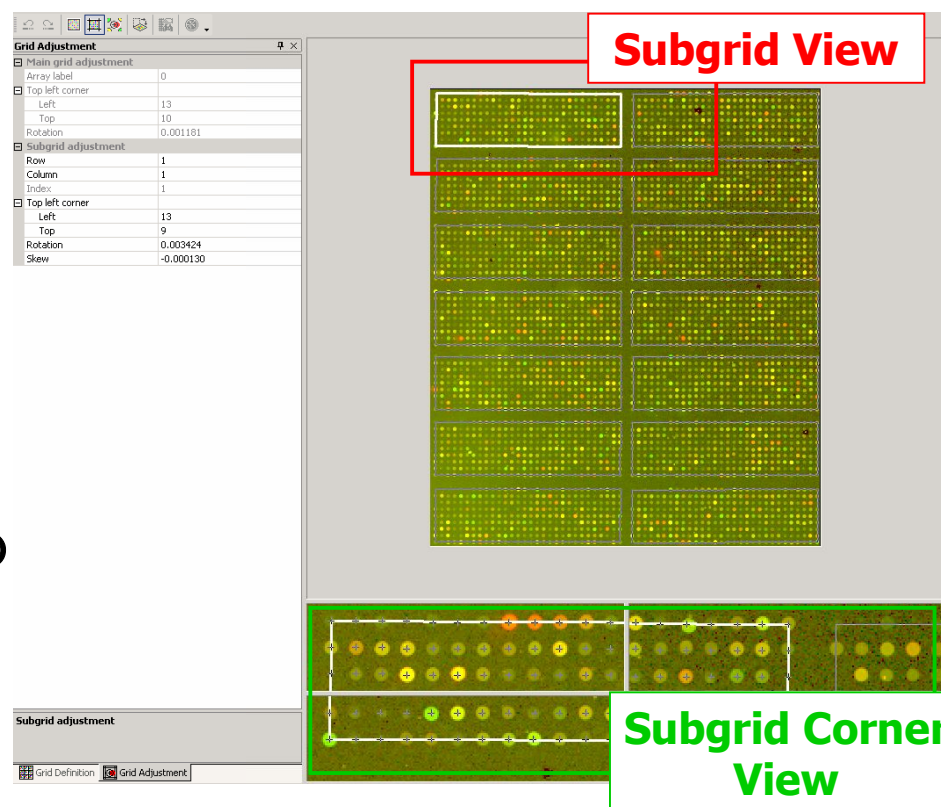
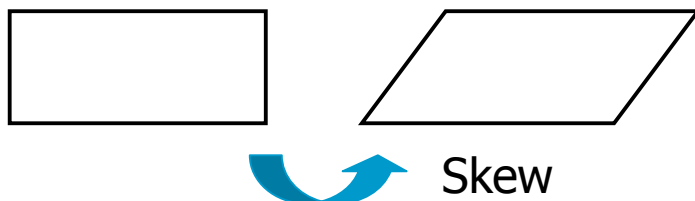
➤ **Subgrid Corner View**

Grid カーソルを使ってグリッドを移動し位置を調整

-  グリッドの角度変更

➤ **Subgrid Corner View**

Skew アイコンをクリックし、アレイの隅にもっていき角度を変更



スポット位置の調整

Adjust Spot



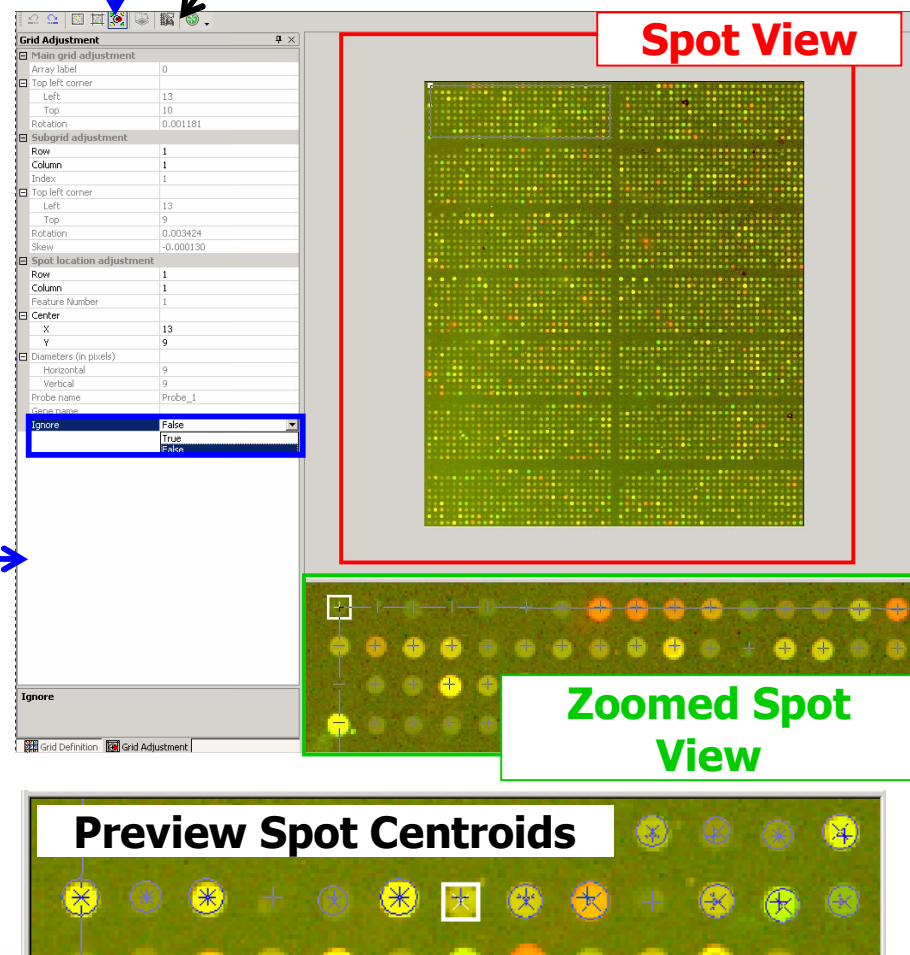
Adjust Spot をクリック

- “+”がスポットの中心にくるように調整
-  をクリックし、**Feature Extraction** がどの位置をスポット中心部として認識するか確認
- スポットを選び、**Grid Adjustment** タブにある **Ignore** をチェックし **True** を選択 

→ このスポットに関して数値化はおこなわれない

Preview Spot Centroids

Adjust Spot



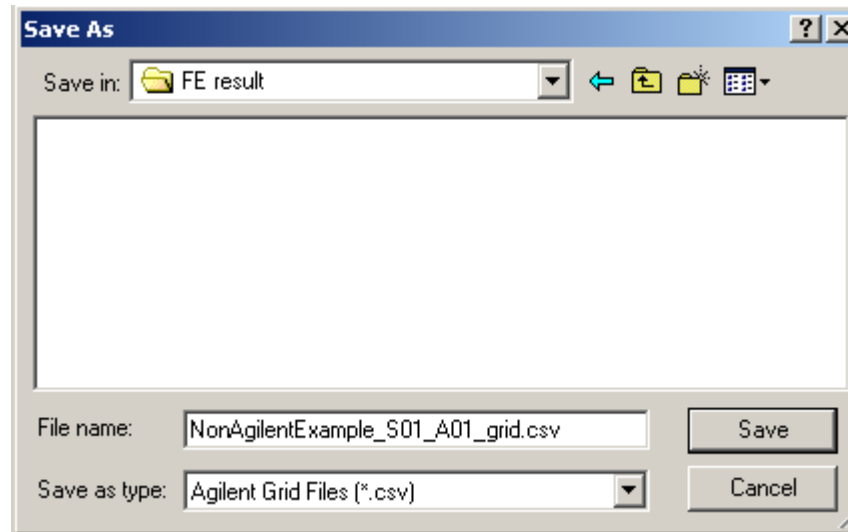
The screenshot shows the software interface with the **Grid Adjustment** tab selected. The **Ignore** checkbox is checked, and the **True** option is selected in the dropdown menu. The **Spot View** is displayed, showing a grid of spots. A **Zoomed Spot View** is also shown, providing a closer look at the spots. The **Preview Spot Centroids** section at the bottom shows a grid of centroids with a central spot highlighted.



グリッドファイルの保存

スポットの中心部が正しく認識されたら Save Grid ボタンをクリックしてファイルをTIFFイメージと同じフォルダに保存

グリッドファイル **_grid.csv** および フィーチャファイル **_feat.csv** が保存される



すぐに数値化を行う場合には、Grid Mode On/Off からグリッドモードを解除し、Feature Extractor を実行

Tab Text File での Gene List 作成: 必要項目

以下のレイアウト、アノテーション情報が必要

- **Grid index (または SubGridRow/SubGridCol)**
 - **SpotRow**
 - **SpotCol**
 - **GeneName または Probename (両方いれることも可能)**
 - **Control 情報も入力可能** -----> オプション
- } 必須項目

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	1	1	1	1	1 FS	Forssman glycolipid synthetase, may be enzymatically inactive; canine homolog FS produces Forss								
2	1	1	1	2	n/a	n/a								
3	1	1	1	3	AMSH	Associated molecule with the SH3 domain of STAM, a protein involved in cytokine signaling that ass								
4	1	1	1	4	n/a	n/a								
5	1	1	1	5	I_1100727	Protein with high similarity to rat CIZ, which is a nucleocytoplasmic shuttling protein that activates m								
6	1	1	1	6	n/a	n/a								
7	1	1	1	7	INAC	Amiloride-sensitive sodium channel (intestine), a member of the DEG/ENaC superfamily of sodium c								
8	1	1	1	8	3-Apr	Protein of unknown function								
9	1	1	1	9	GUCY2F	Guanylyl cyclase F, a membrane-bound guanylyl cyclase that converts GTP to cGMP in the retina p								
10	1	1	1	10	I_932068	Member of the tetraspanin family, has low similarity to kangai 1 (human KA11), which acts as an act								
11	1	1	1	11	FADS1	Fatty acid desaturase 1 (delta 5 desaturase), catalyzes the desaturation of eicosatrienoic acid to ar								
12	1	1	1	12	FLJ20624	Hypothetical protein FLJ20624 (PAK/PLC-interacting protein 1), contains five G protein beta-like WD								
13	1	1	1	13	ST3GALVI	Alpha2-3-sialyltransferase, an enzyme that sialylates glycoproteins and glycolipids, involved in the b								
14	1	1	1	14	AMY1A	Salivary amylase alpha 1A, 1,4-alpha-D-glucan glucanohydrolase which degrades starch to glucose								
15	1	1	1	15	DRG1	Developmentally regulated GTP-binding protein 1 (neural precursor cell expressed developmentally d								



コントロール情報

Control Type

Name	入力コード
Probe	0
Positive	1
Negative	-1
Not Probe*	-20,000
Ignore**	-30,000

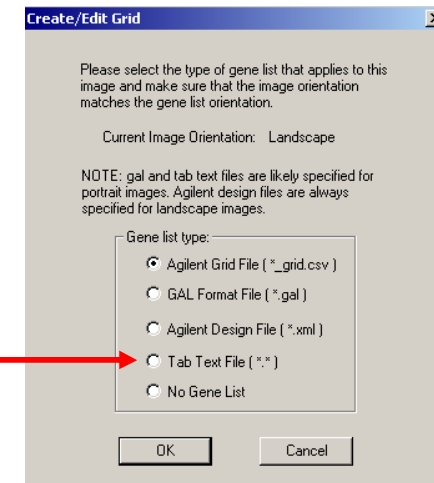
* 数値化され結果に含まれるが、色素バイアスなどの計算には使用されない

** 数値化は行われず、結果にも含まれない



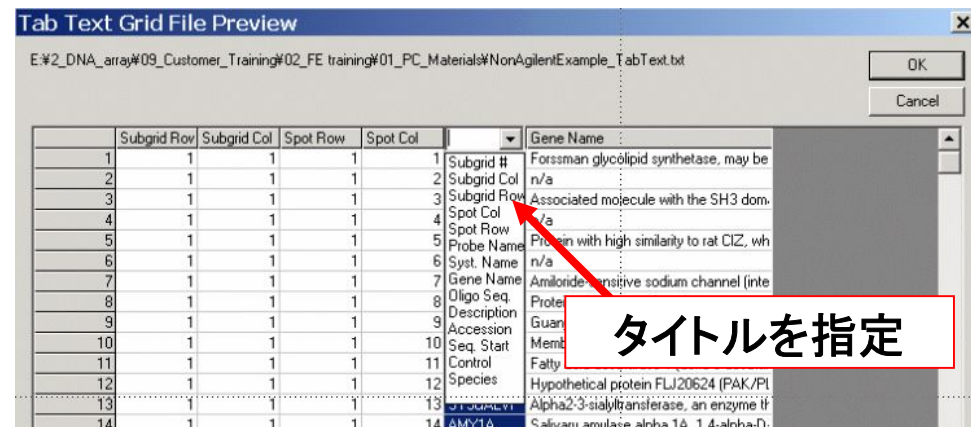
Tab Text File からのグリッド作成

Create/Edit Grid ウィンドウから Tab Text File を選択

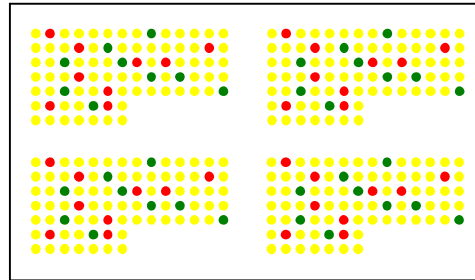


各項目のタイトルを指定して OK をクリック

- **Gene List** の情報をもとにグリッドを自動で作成
- 後のステップは **No Gene List** の時と同様にグリッド位置の調整を行う



角がブランクになっているアレのグリッドあわせ



ブランクの数が少ない場合

Spot location adjustment	
Row	1
Column	1
Feature Number	1
Center	
X	13
Y	9
Diameters (in pixels)	
Horizontal	9
Vertical	9
Probe name	Probe_1
Gene name	
Ignore	False
	True
	False

グリッド作成時に
スポットを1つずつ“Ignore”と
指定（前述）

ブランクの数が多い場合

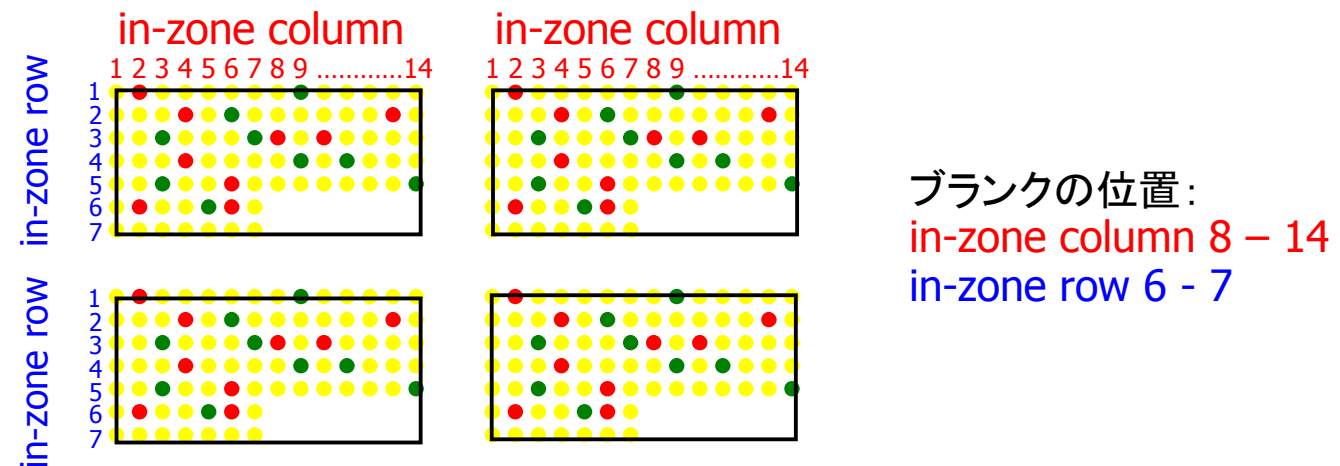
グリッドファイルをまず作成
フィーチャファイルからブランクスポット
を集め “1 (ignore)” と指定

Q	R	S	T	U	
ze	feature nu	no spot	user ignore	control ty	pro
0	1	0	0	0	FS
0	2	0	0	0	n/a
0	3	0	0	0	AMS
0	4	0	0	0	n/a
0	5	0	0	0	I,11
0	6	0	0	0	n/a
0	7	0	0	0	INA

具体的な手順1 ブランクの数が多い場合

フィーチャファイルの編集方法

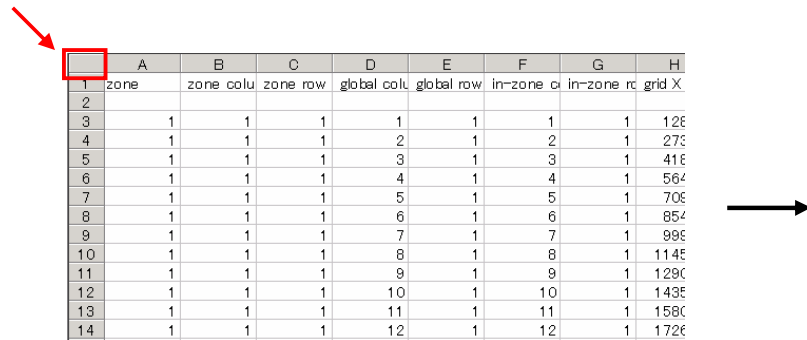
1. No Gene List を選択してグリッドファイルを作成



注意：ブランクの位置もスポットがあると仮定してグリッドファイルが作成

具体的な手順2 ブランクの数が多い場合

2. 作成されたフィーチャファイルを開く

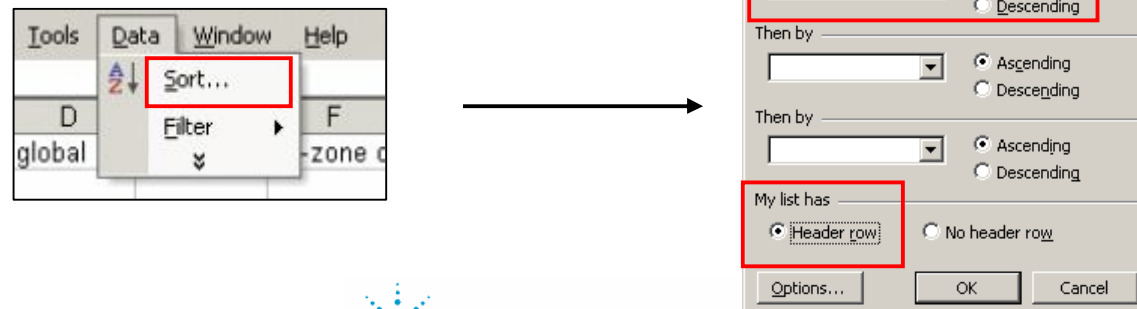


	A	B	C	D	E	F	G	H
1	zone	zone colu	zone row	global colu	global row	in-zone ci	in-zone rc	grid X
2								
3	1	1	1	1	1	1	1	126
4	1	1	1	2	1	2	1	273
5	1	1	1	3	1	3	1	416
6	1	1	1	4	1	4	1	564
7	1	1	1	5	1	5	1	706
8	1	1	1	6	1	6	1	854
9	1	1	1	7	1	7	1	995
10	1	1	1	8	1	8	1	1145
11	1	1	1	9	1	9	1	1290
12	1	1	1	10	1	10	1	1435
13	1	1	1	11	1	11	1	1580
14	1	1	1	12	1	12	1	1726

	A	B	C	D	E	F	G	H
1	zone	zone colu	zone row	global colu	global row	in-zone ci	in-zone rc	grid X
2								
3	1	1	1	1	1	1	1	126
4	1	1	1	2	1	2	1	273
5	1	1	1	3	1	3	1	416
6	1	1	1	4	1	4	1	564
7	1	1	1	5	1	5	1	706
8	1	1	1	6	1	6	1	854
9	1	1	1	7	1	7	1	995
10	1	1	1	8	1	8	1	1145
11	1	1	1	9	1	9	1	1290
12	1	1	1	10	1	10	1	1435
13	1	1	1	11	1	11	1	1580
14	1	1	1	12	1	12	1	1726

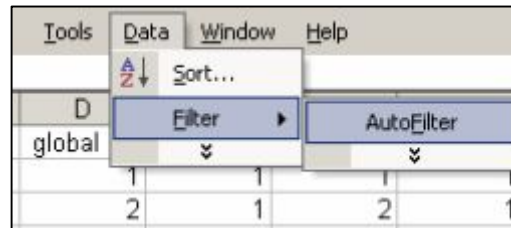
シート左上のブランクのセル(上記図の赤矢印)をクリックして、シート全部を選択

3. メニューバーの **Data>Sort** をクリック **Sort by "zone"** で **Ascending (昇順)** をチェック **My list has "Header row"** にチェック

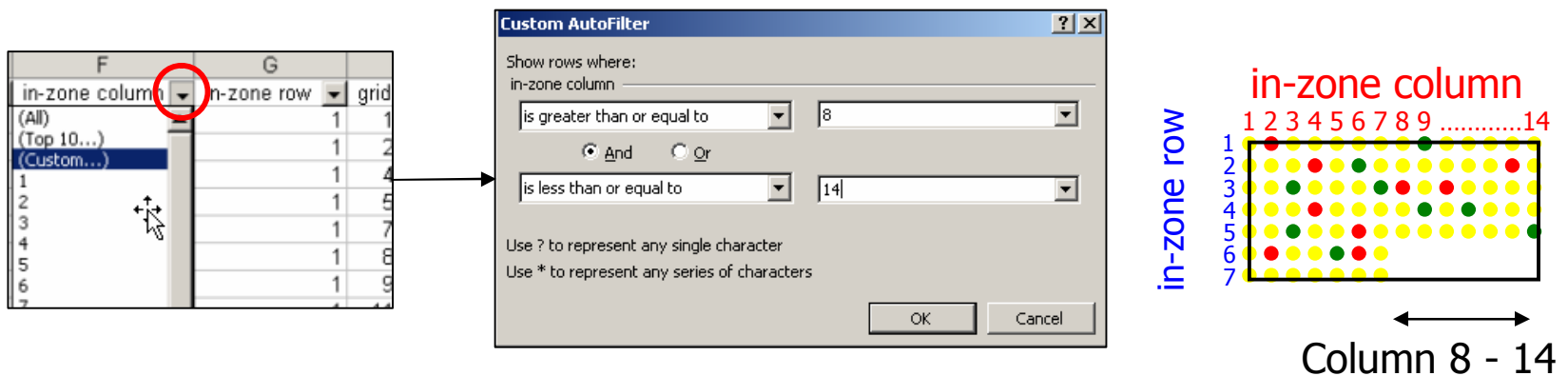


具体的な手順3 ブランクの数が多い場合

4. メニューバーの **Data>Filter>AutoFilter** をクリック

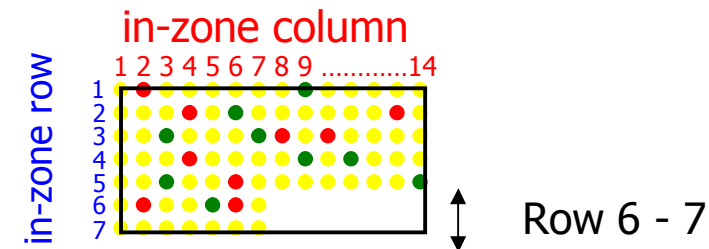


5. “in-zone column” の▼をクリックして、Option を選択 ブランクになっているサブグリッド内の **Column** 番号を入力



具体的な手順4 ブランクの数が多い場合

6. 同様に “in-zone row” の▼をクリックして、Optionを選択
ブランクになっているサブグリッド内の Row 番号を入力



7. フィルターをかけブランクスポットのみを集めた後、
“user ignore” を 0 → 1 に変更

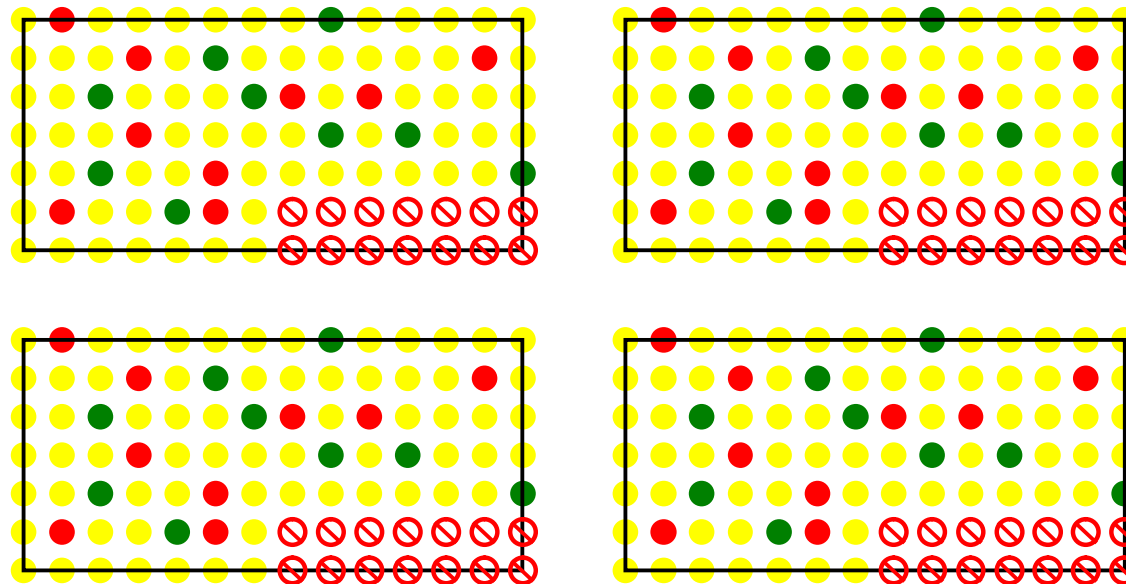
T	
user ignore ▼	
	0
	0
	0
	0
	0
	0
	0

→

T	
user ignore ▼	
	1
	1
	1
	1
	1
	1
	1

具体的な手順5 ブランクの数が多い場合

8. メニューバーの **Data > Filter > AutoFilter** を再度クリック
フィルターを解除後、フィーチャファイルを保存
9. 編集したグリッドファイル、フィーチャファイルを使い **Feature
Extractor** を実行



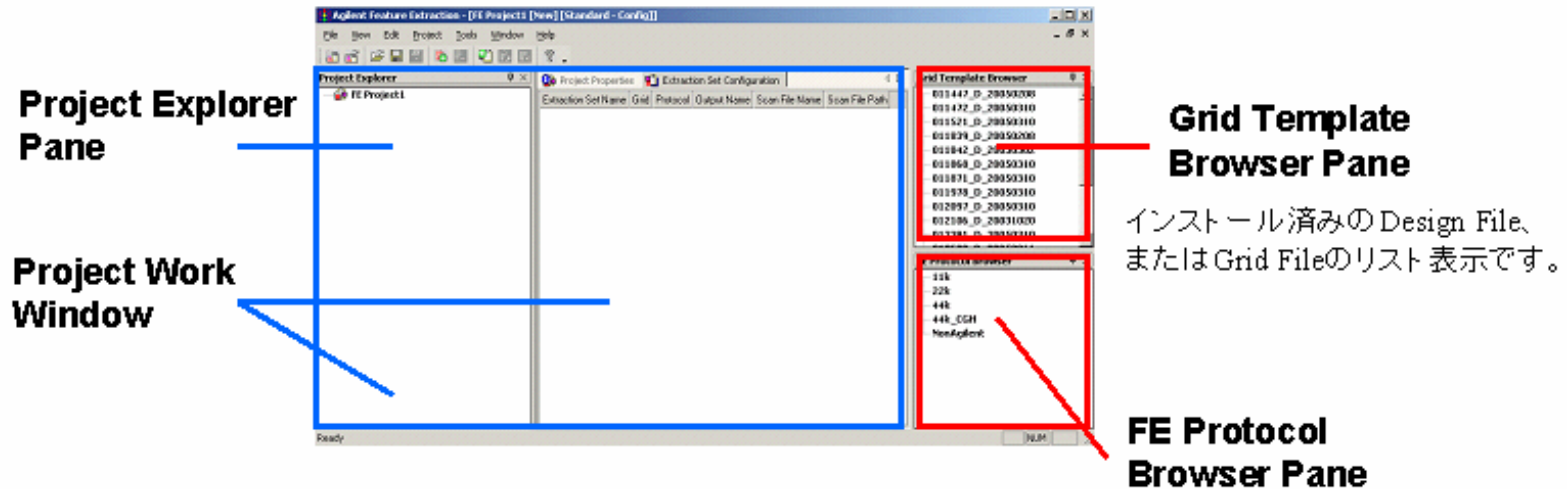
User Ignore が 1 になっているスポット  は数値化結果には含まれません

Feature Extraction

数値化のためのセットアップと 結果について

プロジェクト、グリッドおよびプロトコル

言葉の説明



プロジェクト

- **Project**
1～100枚のTIFFイメージの一括
数値化のプロセス(ファイル保存)
- **Project Properties** および
Extraction Set Configuration
タブ内の設定が必要(後述)

グリッドおよびプロトコル

- **Grid Template**
デザインファイル(アジレントアレイ)
およびグリッドファイル
- **Protocol**
数値化の各ステップのアルゴリズム
の詳細設定

Feature Extraction 数値化に必要なもの

- ① **Agilent** スキャナで読み取った **TIFF** イメージ(クロップの必要なし)
- ② **Grid Template** - デザインファイルまたはグリッドおよびフィーチャファイル

➤ **Agilent** アレイの場合

デザインファイル (.xml)

- カタログアレイキット **CD**
- **Agilent** ウェブから随時ダウンロード可能

➤ **Non-Agilent** アレイの場合

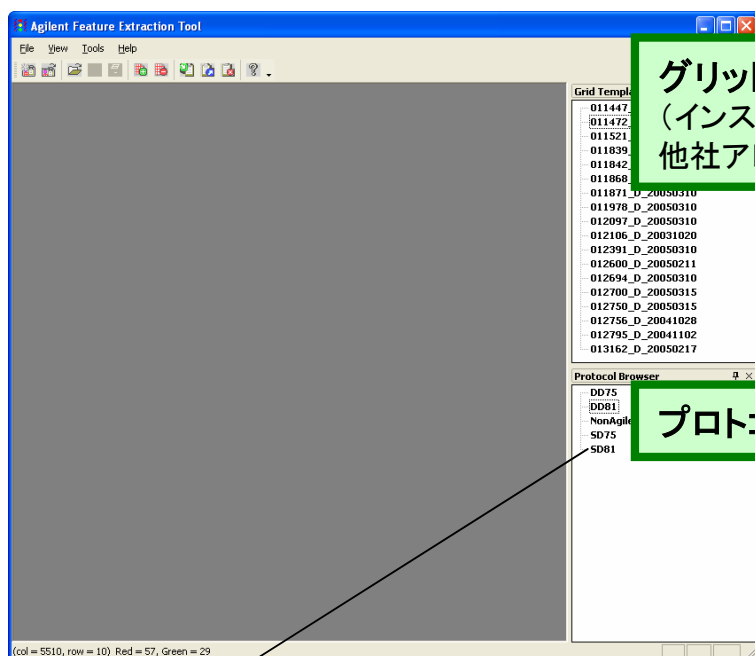
作成したグリッドファイルおよびフィーチャファイル
(**_grid.csv, _feat.csv**)

- ③ **Protocol** - 数値化のためのプロトコル



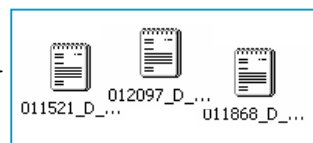
グリッドテンプレートおよびプロトコルの変更・追加

ダブルクリックにて確認・変更 & 右クリックで追加



グリッドテンプレート
(インストール済みのデザインファイルや、
他社アレイのデザインファイルを格納)

追加



新デザインファイル
(アジレントアレイ)

追加

44k CGH 0605

新プロトコル

Agilent
website

Microarray Kits webpage

[Technical Support](#)

- [Search Gene List](#)
- [Download Gene List](#)
- [Download Design Files](#)

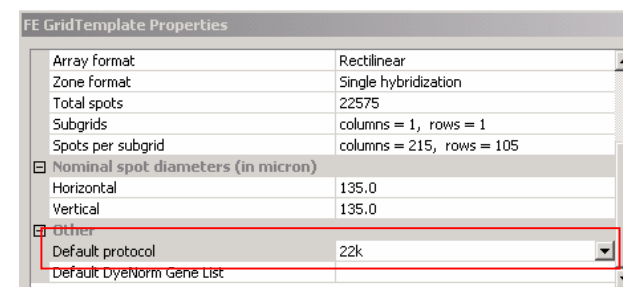
Feature Extraction
Software webpage

[Technical Support](#)

- [Frequently Asked Questions](#)
- [Specifications](#)
- [Download Software](#)
- [Download Trial License](#)
- [new Download Protocols](#)

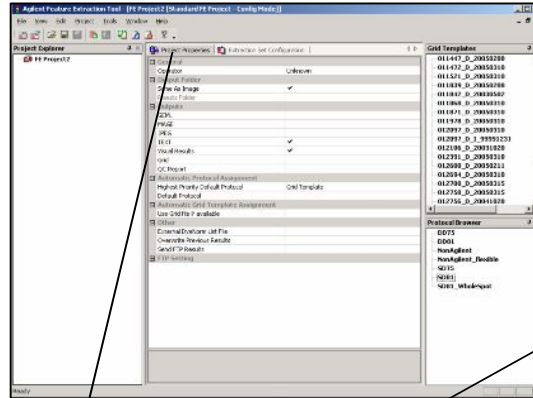
既存の“プロトコル”のパラメータを
変更して別の名前で保存すること
ができます。ダブルクリックしてパ
ラメータを変更後、“Save as...”で
保存すると、Protocol Browserに追
加されます。

新“プロトコル”を追加したり既存のプロ
トコルを変更して名前を変えて保存した場
合、対応するデザイン(グリッド)ファイル
をダブルクリックして、“Properties”(右
図)の”Default Protocol”を変更すること
で、新プロトコルとデザインファイルを自
動マッチングできるようになります。

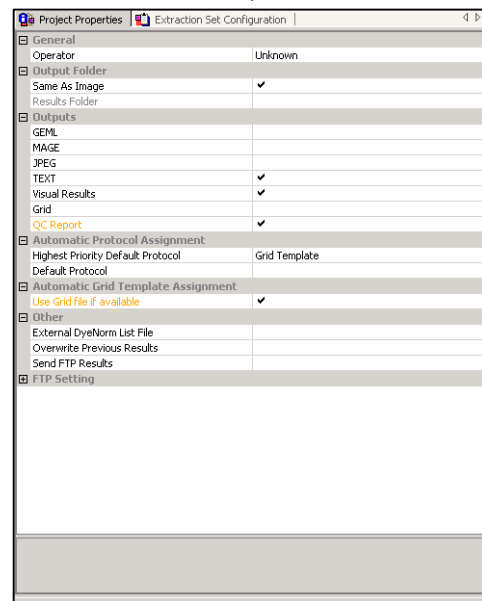


プロジェクトファイルの作成

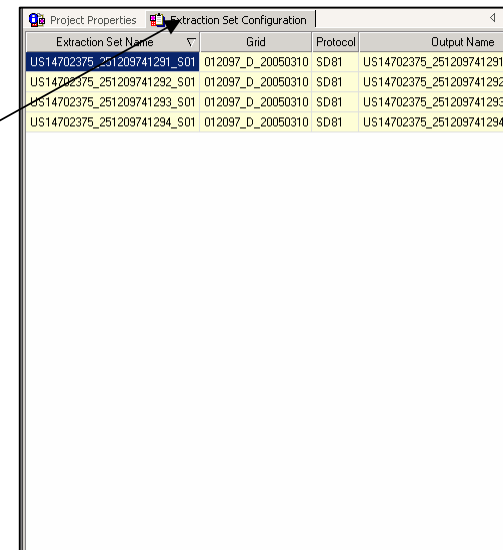
(1) File>New>Standard Project を開く(新規作成)



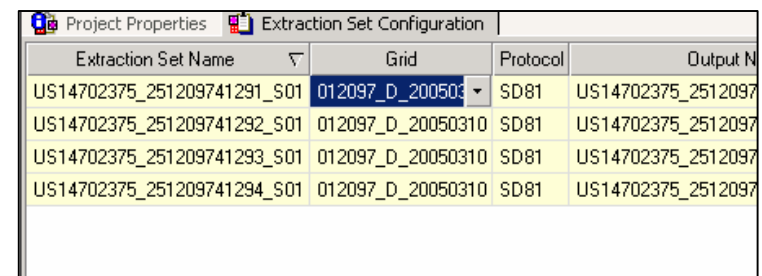
(2) Project Propertiesを設定



(3) Extraction set configurationタブを開いて、ここにtiffをドラッグ



(4) GridおよびProtocolにはバーコード番号を元に自動で読み込まれた設定が入るが、タブで選択変更可能



Feature Extraction の主なアウトプットファイル

- **TEXT**

タブ切テキストフォーマット（Excel やGeneSpringなどで読み込み可能）

- **Geneview(miRNAのみ)**

miRNA数値化結果（Excel やGeneSpringなどで読み込み可能）

- **JPEG**

イメージ圧縮ファイル（注意：JPEG ファイルからスポット数値化不可能）

- **MAGE**

MicroArray Gene Expression Markup Language フォーマット

- **Visual Result**

TIFFイメージに重ねてフラグ情報等確認できるファイル

- **QC Report**

1枚のアレイの数値化結果のレポート（統計情報およびプロット）



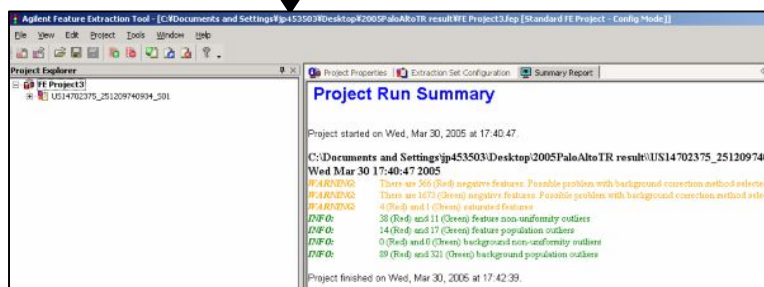
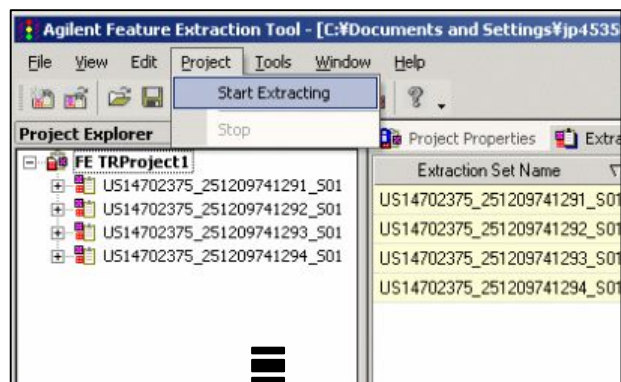
MAGE について

- XML言語をベースとして開発
- **Rosetta** 社の **Resolver** ソフトウェアに インポート 可能
- マイクロアレイに関する情報の詳細の記述、共有のためにデザインされた言語
- マイクロアレイデザイン、製造情報、実験設定、実験操作情報、遺伝子発現データ、データ解析結果の詳細を記述可能
- **ArrayExpress (EBI)** , **GEO (NCBI)** や **CIBEX (DDBJ)** などの公的マイクロアレイデータベースで採用されているフォーマット



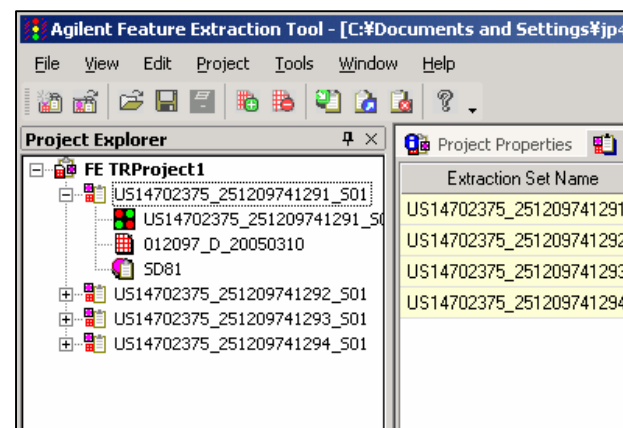
プロジェクトファイルのRun (バッチ処理の開始)

(5)Project>Start Extraction で開始



Runが終わると、Summary reporterというタブができて、情報が表示されます。

(6)Extractionが終わったら。Project Explorerのツリービューのなかからマークのファイルをダブルクリックすると画像の画面へ。



Resolver へのファイル転送

FTP 転送ファイル

- **MAGE**
- **TIFF** あるいは **JPEG** 画像

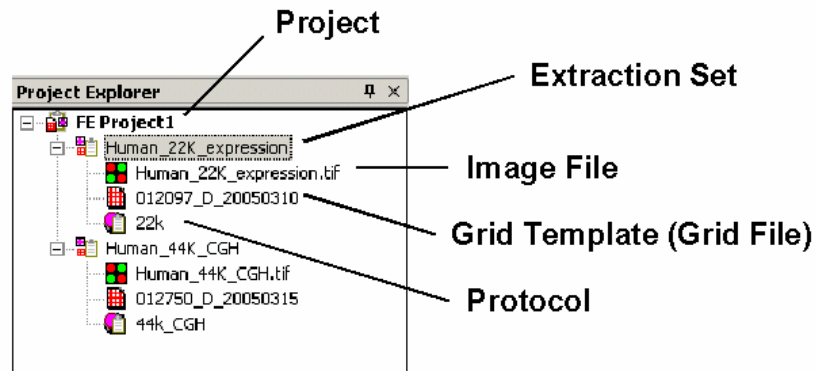
FTP 設定

- **Destination**
- **FTP port**
- **User name**
- **Password**

Send FTP Results	True
FTP Setting	
Result Type	MAGE
Image Type	JPEG
Machine Destination	
Port	21
User Name	resolverftp
Password	*****
Pattern Destination Folder	mage
GEML/MAGE Destination Folder	mage



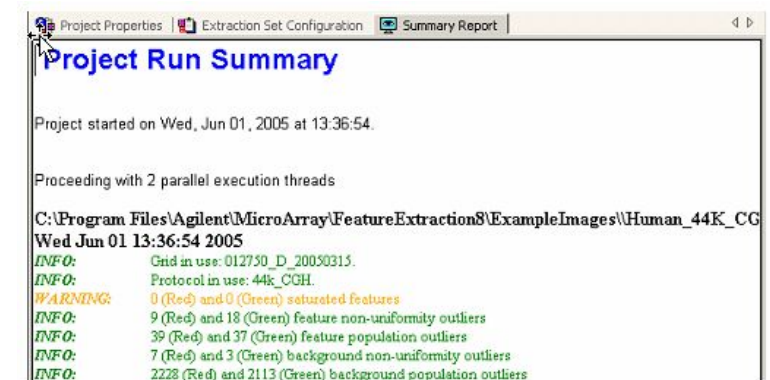
プロジェクト進行の確認



Running Monitor



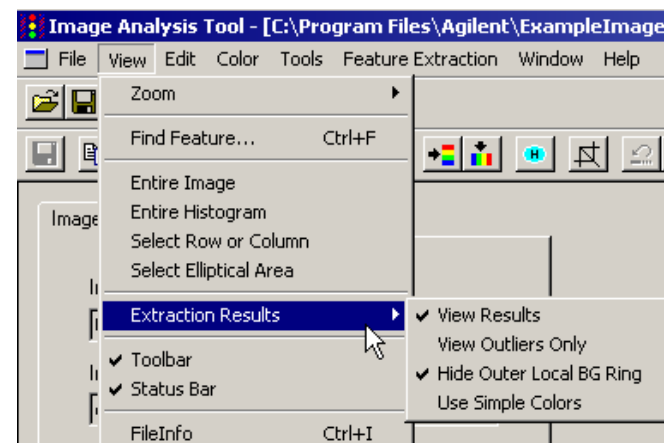
FE Project name_
YYYYMMDDHHMM.rtf



Visual Results でのフラグ確認

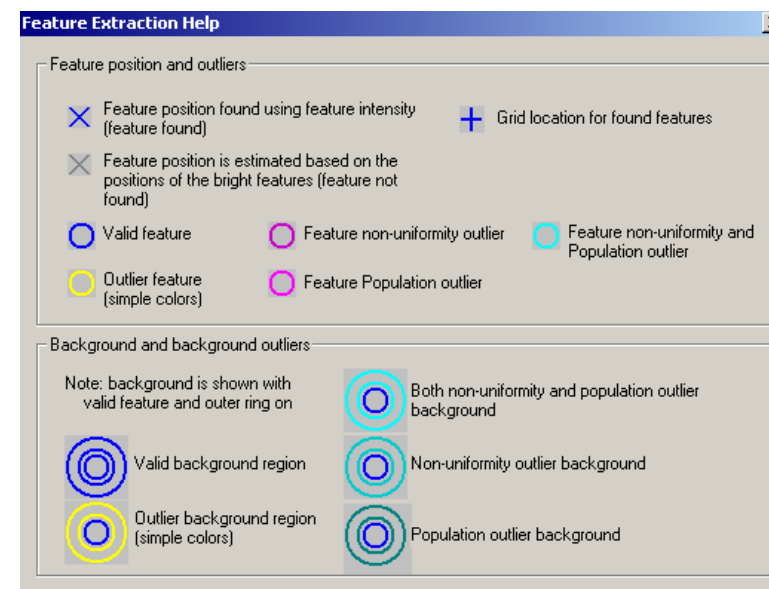
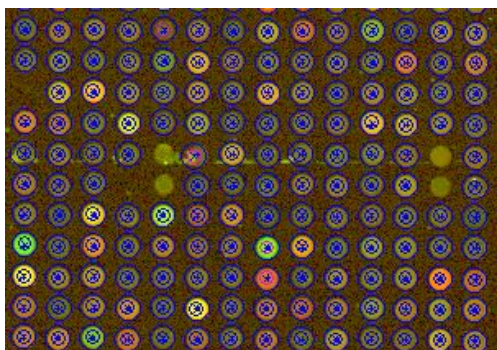
View > Extraction Results

- **View Results**
- **View Outlier Only**
- **Hide Outer Local BG Ring**
- **Use Simple Colors**



Help > Feature Extraction Output Quick Reference

- フラグ内容の確認



Feature Extraction テキスト結果例

設定値

TYPE	text	text	text	text	text	integer	integer	float	float	float	float	float	text	float	float	float
FEPARAM	FeatureEx	FeatureEx	FeatureEx	SpotFinder	SpotFinder	SpotFinder	SpotFinder	SpotFinder	SpotFinder	SpotFinder	SpotFinder	SpotFinder	CornerMet	SpotFinder	SpotFinder	CornerUL_x
DATA	uytruong	UYTRUON	Wednesda	14f32760-1	Version A.	105	215	5	25.00001	21.47222	6.75	6.75	Auto Find	0.3	0.5	101.31925

統計値

TYPE	float	float	float	integer	float	float	float	integer	float	float	integer	float	float	float	integer	integer
STATS	gDarkOffse	gDarkOffse	gDarkOffse	gDarkOffse	rDarkOffse	rDarkOffse	rDarkOffse	rDarkOffse	gNumSatF	gLocalBGI	gLocalBGI	gLocalBGI	gGlobalBGI	gGlobalBGI	gGlobalBGI	rNumSatFe
DATA	353.968	354	4.52951	1000	326.662	326	10.5443	1000	0	375.762	0.602773	21111	375.762	0.602773	21111	30

データ

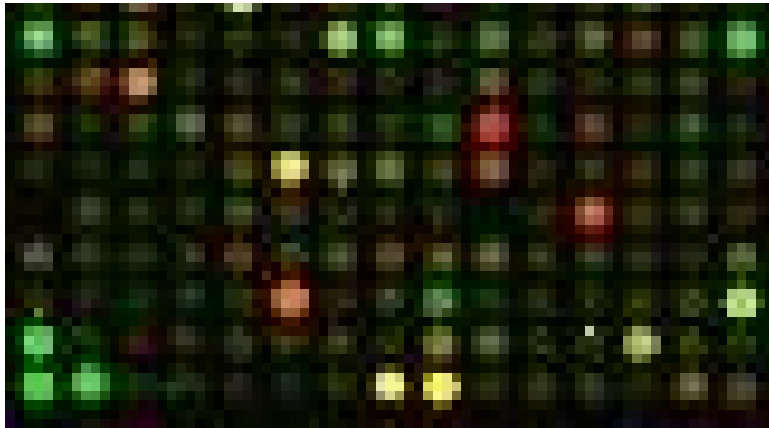
TYPE	integer	integer	integer	text	text	text	text	integer	integer	text	text	text	text	float	float	float
FEATURE	FeatureNu	Row	Col	ug	mg	LocusLink	gb	ProbeUID	ControlTyp	ProbeNam	GeneNam	Systemati	Description	PositionX	PositionY	LogRatio
DATA	1	1	1					0	1	(+)Pro25G	Pro25G	Pro25G		101.385	87.6783	-2.00E+00
DATA	2	1	2					1	-1	(-)3xSLv1	NegativeCr	NegativeControl		122.204	87.4314	2.24E-01
DATA	3	1	3	Mm.2287	1347009	26442	NM_01196	2	0	A_65_P00	Psma5	L0267E07	proteasom	143.691	87.9431	9.27E-02
DATA	4	1	4	Mm.19945	101900	17387	NM_00860	3	0	A_65_P00	Mmp14	U549B4.1	matrix met	165.472	86.8924	-3.62E-01
DATA	5	1	5	Mm.1571	99217	12552	NM_00986	4	0	A_65_P00	Cdh11	H3149F09	cadherin 1	186.246	87.9051	-1.13E+00
DATA	6	1	6	Mm.25353		66527		5	0	A_65_P00	22104170	C0152H05	RIKEN cDI	207.58	87.9296	3.90E-01
DATA	7	1	7					6	1	(+)Pro25G	Pro25G	Pro25G		229.961	87.8839	-2.00E+00
DATA	8	1	8	Mm.221899				7	0	A_65_P01	L0951F09	L0951F09	ESTs	250.832	87.9961	-2.52E-02
DATA	9	1	9	Mm.173571				8	0	A_65_P01	H3065B10	H3065B10	ESTs, We	271.327	88.1006	-9.86E-02
DATA	10	1	10					9	0	A_65_P01	C0267B04	C0267B04	NA	292.344	88.2031	-7.75E-02
DATA	11	1	11	Mm.4557	1346023	24086	AK014829	10	0	A_65_P01	Ttk2	H3103D06	tousled-lik	313.6	87.8859	6.40E-02
DATA	12	1	12	Mm.56287	1340029	17771	NM_01084	11	0	A_65_P01	Mti5	H3158G12	metallothi	334.645	88.2625	-5.90E-02
DATA	13	1	13					12	0	A_65_P01	J1002D05	J1002D05	NA	356.082	88.7195	-2.37E-02
DATA	14	1	14					6	1	(+)Pro25G	Pro25G	Pro25G		375.222	88.9569	-2.00E+00
DATA	15	1	15	Mm.207090				13	0	A_65_P01	L0253B01	L0253B01	ESTs	397.126	89.8382	-7.12E-04
DATA	16	1	16	Mm.222562				14	0	A_65_P01	C0162E08	C0162E08	ESTs, Higl	418.386	89.7211	1.37E-01
DATA	17	1	17	Mm.217306				15	0	A_65_P02	K0515H11	K0515H11	ESTs	439.11	90.1722	-3.43E-02
DATA	18	1	18	Mm.50610		215114		16	0	A_65_P02	MGC27616	K0310D04	hypothetic	460.634	89.1909	-5.88E-01
DATA	19	1	19	Mm.133186		214133	NM_14598	17	0	A_65_P02	MGC37386	H3046H03	hypothetic	481.957	89.953	-1.88E-01
DATA	20	1	20	Mm.172782		97969		18	0	A_65_P02	C79601	J0068H05	expressed	502.844	90.0417	-9.79E-02
DATA	21	1	21					6	1	(+)Pro25G	Pro25G	Pro25G		524.333	89.9036	-2.00E+00
DATA	22	1	22	Mm.216206		105648		19	0	A_65_P02	AWD47562	K0345D09	expressed	545.652	90.0507	2.24E-02
DATA	23	1	23					20	0	A_65_P02	J0460H10	J0460H10	NA	566.969	90.6743	-6.94E-03
DATA	24	1	24	Mm.172684		97144	XM_16100	21	0	A_65_P02	C78128	H3039C07	expressed	588.065	90.3663	6.84E-02
DATA	25	1	25	Mm.42027	108426	16561		22	0	A_65_P02	Kif1b	L0517F11	kinesin fan	609.346	90.8823	-4.87E-02
DATA	26	1	26			232798	XM_13313	23	0	A_65_P02	I9330175N	K0322H09	hypothetic	630.52	90.8456	2.77E-02
DATA	27	1	27	Mm.38450	1858222	53860	NM_01738	24	0	A_65_P02	9-Sep	L0226E02	septin 9	651.296	91.2936	3.20E-01
DATA	28	1	28					6	1	(+)Pro25G	Pro25G	Pro25G		673.73	90.8139	-2.00E+00
DATA	29	1	29					25	0	A_65_P03	C0401E10	C0401E10	NA	693.949	91.2498	9.66E-02

※Output Package設定で、FullおよびCompactを選択可能

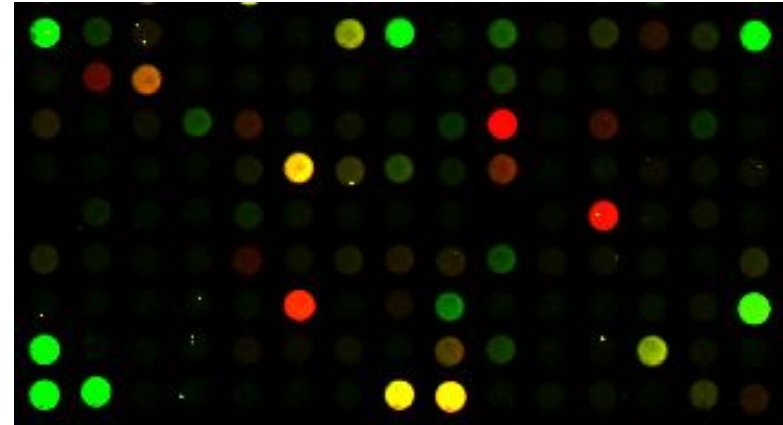


JPEGに関する注意

JPEG



TIFF



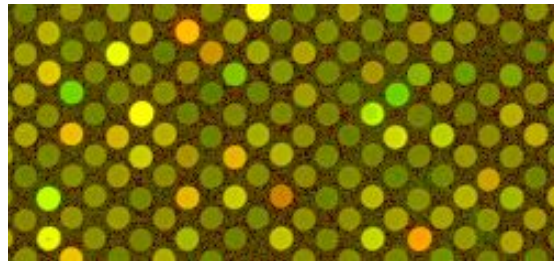
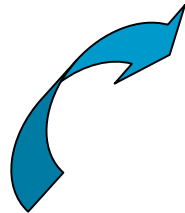
JPEG → TIFF と比較して、解像度が低い

JPEG から Feature Extraction (スポット数値化)は実行できない

※TIFF は必ず保存しておく

データワークフロー

Grid Template
Grid File
FE Protocol



TIFF画像

Agilentスキャナで読み込んだ
Agilentおよびnon-Agilentマイクロアレイ

GenePixスキャナ*で読み込んだ
Agilentマイクロアレイ

*Gene Pix 4000Bスキャナで読み込んだ
TIFF画像に制限



Feature Extraction
(数値化ソフトウェア)



GeneSpring

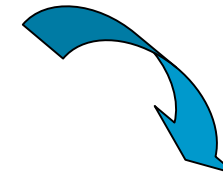


解析ソフトウェア



Agilent Technologies

スポット検出
フィーチャ&バックグラウンド領域の決定
シグナル強度の抽出
バックグラウンド補正
色素補正
発現比計算



QC Report
Grid
Shape
Text
JPEG
MAGE-ML
GEML

Import

FTP
転送

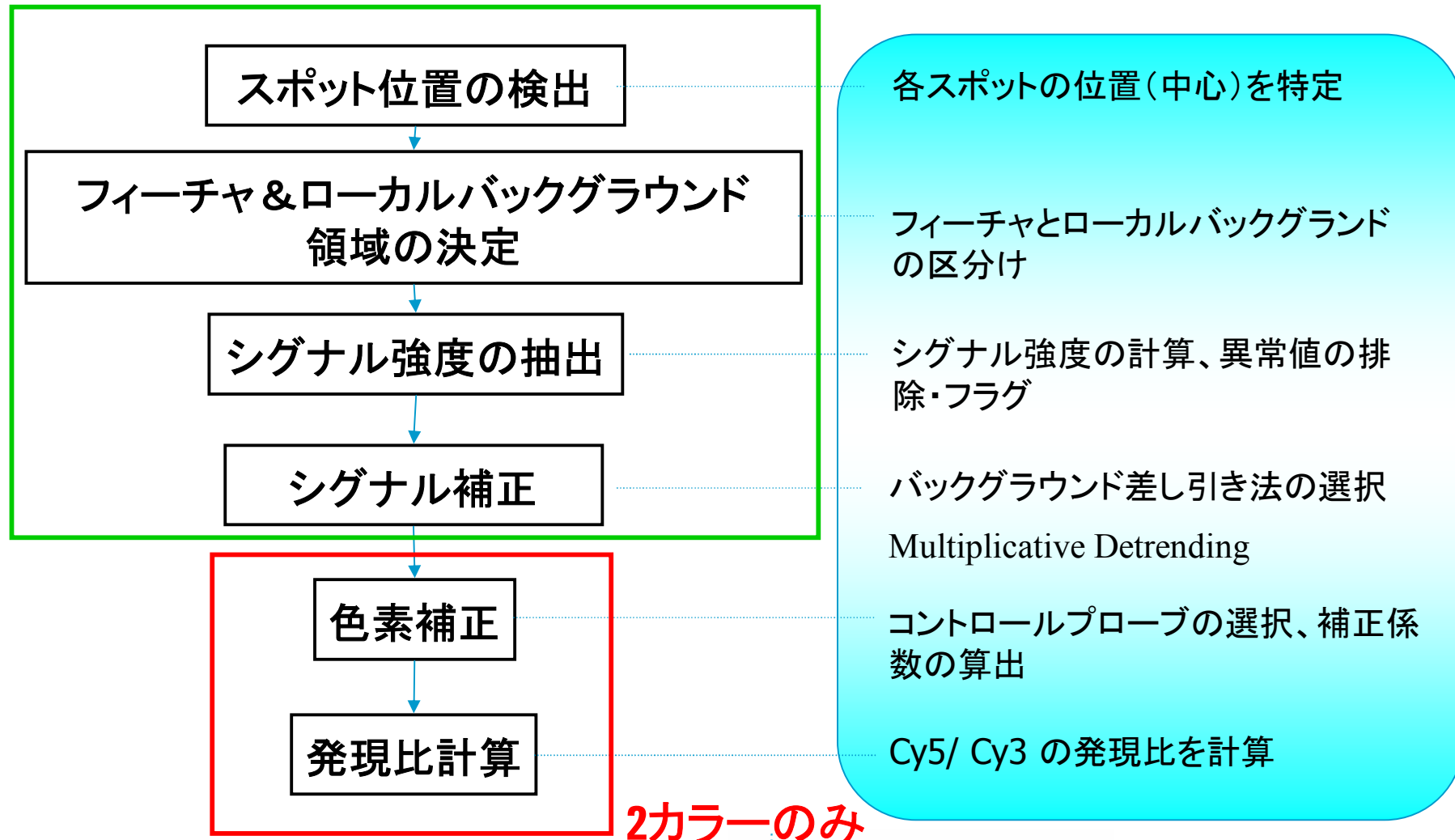
Feature Extraction
アウトプットファイル

Feature Extraction アルゴリズム



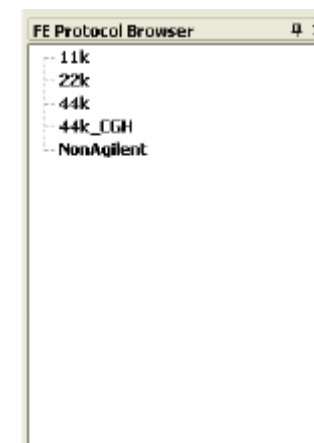
数値化(スポット解析)の流れ

1、2カラー共通

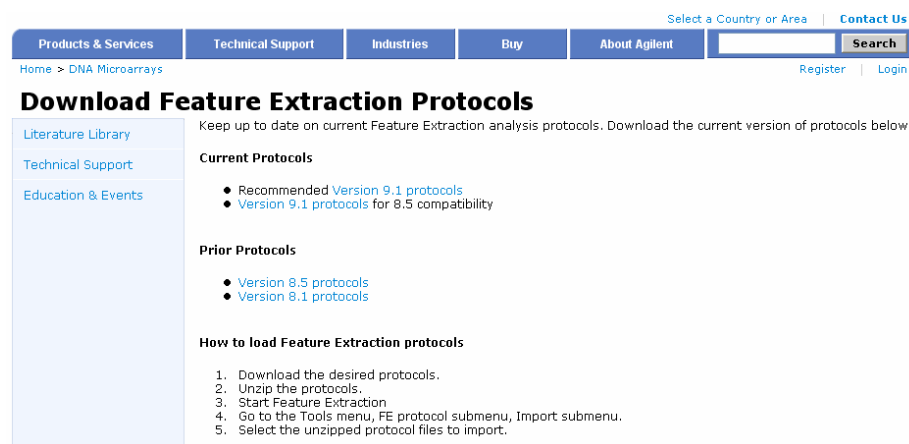


Feature Extraction デフォルトのプロトコル

- **Agilent システムを使用した際の推奨**
 - アプリケーションによって設定は異なる
(**Gene Expression, CGH, miRNA**)
 - **Gene Expression**用(1色用、2色用)
- **Non-Agilent:** アレイ・システムの場合、マイクロアレイの種類やラベル化等の違いを考慮した設定が必要



Feature Extractionプロトコルファイルのダウンロードサイト

A screenshot of the Agilent website's "Download Feature Extraction Protocols" page. The page has a navigation bar with links like "Products & Services", "Technical Support", "Industries", "Buy", "About Agilent", and a search bar. Below the navigation bar, the page title is "Download Feature Extraction Protocols". The main content area includes a sub-header "Keep up to date on current Feature Extraction analysis protocols. Download the current version of protocols below." and two sections: "Current Protocols" and "Prior Protocols". The "Current Protocols" section lists "Recommended Version 9.1 protocols" and "Version 9.1 protocols for 8.5 compatibility". The "Prior Protocols" section lists "Version 8.5 protocols" and "Version 8.1 protocols". At the bottom, there is a section titled "How to load Feature Extraction protocols" with a numbered list of five steps: 1. Download the desired protocols. 2. Unzip the protocols. 3. Start Feature Extraction. 4. Go to the Tools menu, FE protocol submenu, Import submenu. 5. Select the unzipped protocol files to import.

<http://www.chem.agilent.com/scripts/generic.asp?lpage=35952&indcol=Y&prodcol=Y>



現在、Webからダウンロード可能なプロトコル (2007年4月現在)



Agilent Technologies

Our measure is

Select a Country or

Products & Services

Technical Support

Industries

Buy

About Agilent

Home > DNA Microarrays

Download Feature Extraction Protocols

Literature Library

Technical Support

Education & Events

Keep up to date on current Feature Extraction protocols. Download the current version of pro

Current Protocols - Version 9.5

※各Protocolの詳細は、Protocol Useに記載

- **new** Complete set of protocols (83 KBytes)

Protocol Use

Complete set of protocols now include:

- CGH-v4_95_Feb07 (11 KBytes)
- CGH-v4_95_Tecan_Feb07 (11 KBytes)
- GE1-v5_95_Feb07 (10 KBytes)
- GE1-v1_95_Feb07 (10 KBytes)
- GE2-v5_95_Feb07 (11 KBytes)
- GE2-v4_95_Feb07 (11 KBytes)
- GE2-SSPE_95_Feb07 (11 KBytes)
- GE2-NonAT_95_Feb07 (10 KBytes)

Current Protocols – Version 9.5

• Complete set of protocols

Prior Protocols – Version 9.1

• Complete set of protocols



初期設定を調べるためには？

- *Reference Guide* 参照

例) 遺伝子発現1色法の プロトコル



Agilent Feature Extraction Software
Reference Guide

1 Default Protocol Settings

Default protocol settings—an introduction 12
Tables of Default Protocol Settings 14
Differences in protocol settings based on each step 45

See Chapter 4, “Changing Protocol Settings” in the User Guide to learn the purpose of all the parameters and settings and how to modify them.

When the software assigns a protocol to an extraction set, the software loads a set of protocol parameter values and settings that affect the process and results for Feature Extraction.

Parameter values in the protocol depend on the microarray type and your experiment. The following pages list the default settings for each of the protocol templates shipped with the software. Each protocol template represents a different microarray type. You can view these settings and values when you open the Protocol Editor for each of the protocol templates.

1 Default Protocol Settings Tables of Default Protocol Settings

Tables of Default Protocol Settings

GE1-v1_95 protocol

This is a 1-color gene expression protocol for use with the Gene Expression Analysis lab protocol (version 1-publication number G4140-90040).

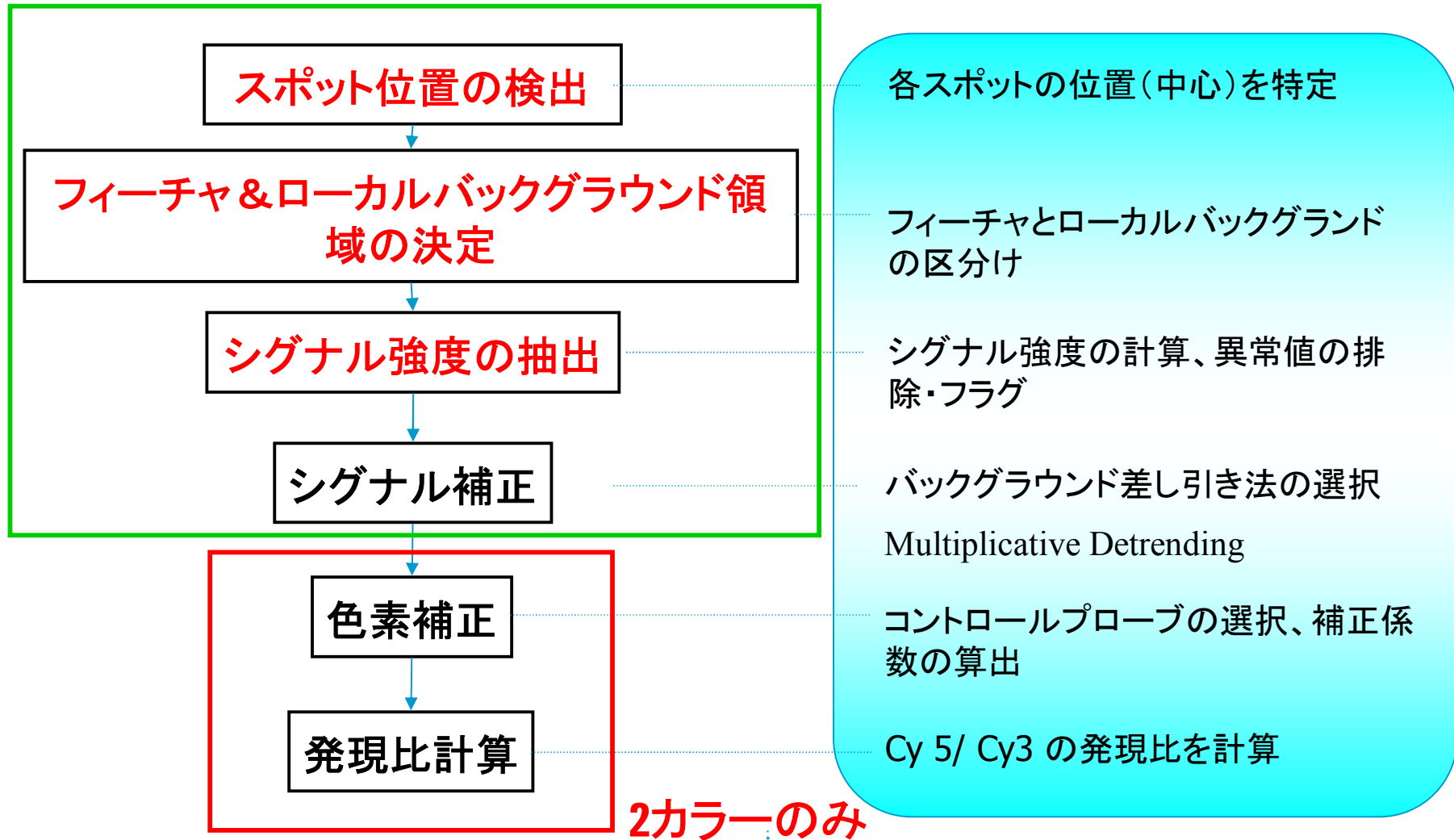
Table 2 Default settings for GE1-v1_95 protocol

Protocol Step	Parameter		Default Setting /Value (v.0.5)
Place Grid	Array Format	For any format automatically determined or selected by you, the software uses the default Placement Method listed below.	Automatically Determine [Recognized formats: Single Density (11k, 22k), 25k, Double Density (44k), 95k, 185k (5 and 10 u), 244k (5 and 10 u) and Third Party]
	Placement Method	The parameters and values for placing the grid differ depending on the format, but you can't see the differences because the values are hidden.	Allow Some Distortion
Find Spots	Spot Format	Depending on the format selected by the software or by you, the default settings for this step change. See the rows below for the default values for finding spots.	Automatically Determine [Recognized formats: same as those listed above]
		Use the Nominal Diameter from the Grid Template	True (All Formats)



数値化(スポット解析)の流れ

1、2カラー共通

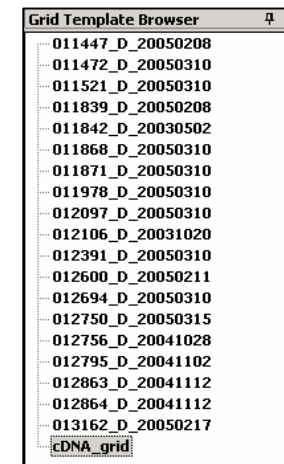
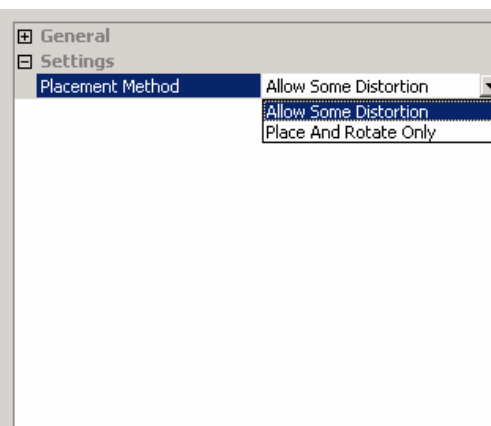
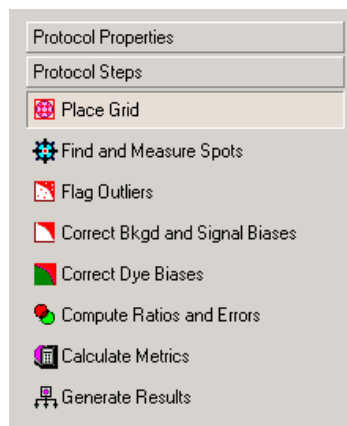




Place Grid グリッドの初期化（自動）

1. グリッドの呼び出し

グリッドテンプレート
(デザインファイル)
を使い自動認識



2. Placement Method(配置方法):

- **Allow Some Distortion (Default)**
フィーチャ／サブグリッド間の間隔
および角度変更を許可
- **Place And Rotate Only**
グリッドの回転位置のみを許可

注意: FE v8 よりアジレントアレイのグリッドの初期化アルゴリズムが変更になりました。TIFFイメージをクロップする必要はありません。アルゴリズムの変更に伴い、スポット中心位置が微妙に異なる場合があります(次頁参照)。



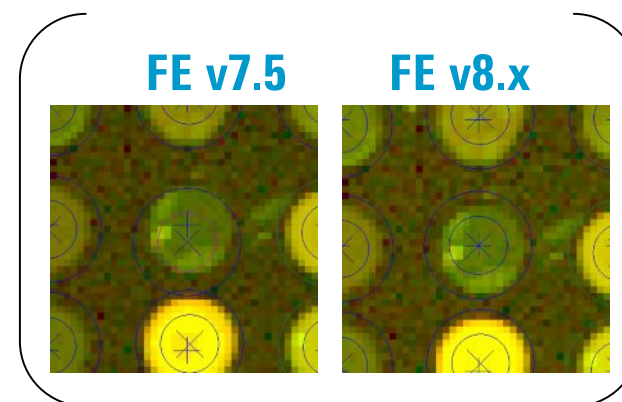
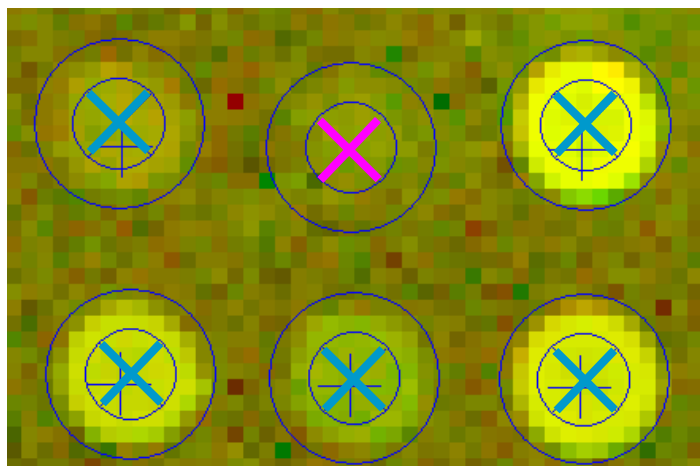
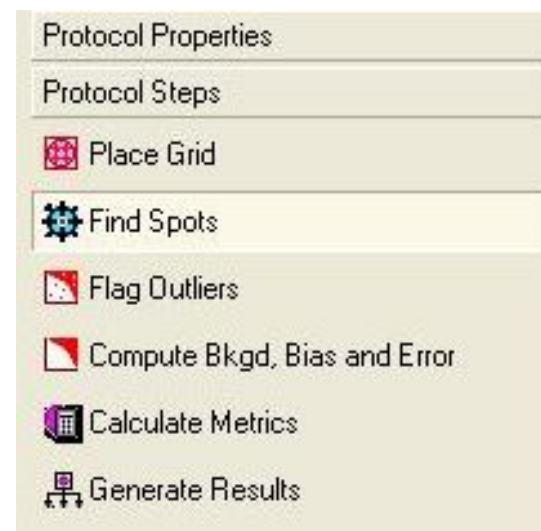


Find and Measure Spots

アルゴリズムの流れ

1. 全てのスポット中心位置の決定

- 明るいスポット 
グリッド情報から微調整して
スポット中心位置を決定
- 暗いスポット 
グリッド位置情報から中心位置を決定

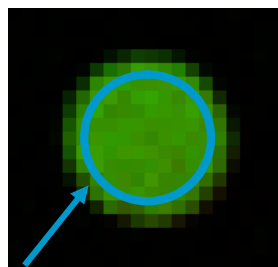




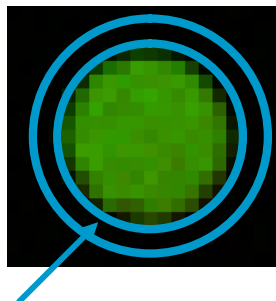
Find and Measure Spots

アルゴリズムの流れ(続き)

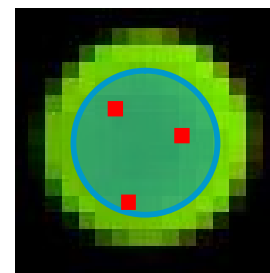
フィーチャ領域決定



ローカルバックグラウンド
領域決定



異常ピクセル排除



■ 外れ値を持つピクセル



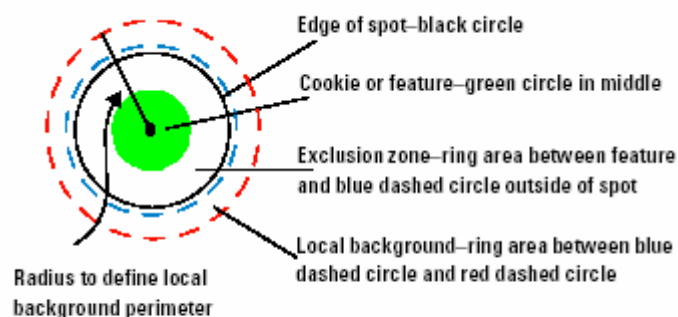


Find and Measure Spots

フィーチャとバックグラウンド領域の定義

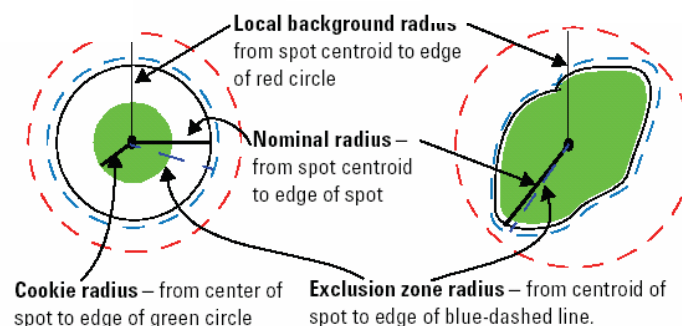
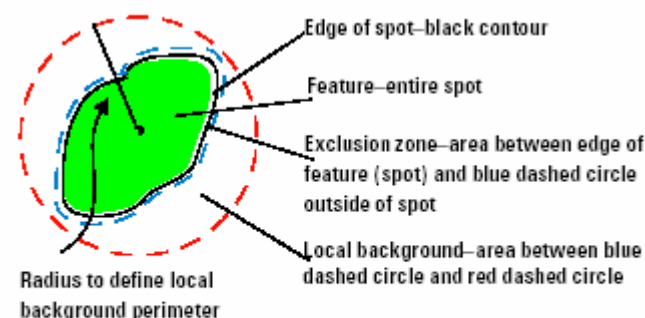
Cookie Cutter 法

スポット中心部から円を描き
フィーチャ領域を決定



Whole Spot 法

スポット形状に合わせて
フィーチャ領域を決定



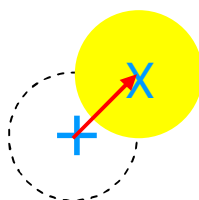
径の定義





Find and Measure Spots **Spot Deviation Limit**

- グリッド位置からのずれの許容範囲
 - 初期設定値: **1.5 (Agilent Oligo Microarray)**
 - ユーザー設定: 名目径(半径)に乗ずる値を設定
- **Dev Limit を小さくする**
スポット中心部の可動範囲が狭くなるので、スポットが極端にずれている場合、スポットが見つけにくくなる
- **Dev Limit を大きくする**
近接しているスポットあるいはごみを誤認識する可能性がある



- X 実際のスポットシグナルから修正後の中心位置
- Dev Limit によるずれの許容範囲
- + グリッド情報からの推定したスポット位置





Find and Measure Spots

フィーチャとローカルバックグラウンド領域の設定

Protocol Properties

Protocol Steps

Place Grid

Find Spots

Flag Outliers

Compute Bkgd, Bias and Error

Correct Dye Biases

Compute Ratios

Calculate Metrics

Generate Results

General

Settings

Spot Format	ThirdParty
Use The Nominal Diameter From GridTemplate?	True
Spot Deviation Limit	1.50
Calculation of Spot Statistics Method	Use Cookie
Cookie Percentage	0.650
Exclusion Zone Percentage	1.200
Auto Estimate the Local Radius	True
Pixel Outlier Rejection Method	Inter Quartile Region
RejectIQRFeat	1.42
RejectIQRBG	1.42
Statistical Method for Spot Value from Pixels	Use Mean / Standard Deviation

名目径をグリッドテンプレート(デザインファイル)
または
グリッド配置アルゴリズムから取得

Cookie またはWhole Spot を選択

Use Cookie
Use Whole Spot

- **Cookie Cutter** を選択した場合、**Cookie Percentage** を設定可能
- **Exclusion Zone Percentage** を設定可能 - 初期設定値: **1.2** (アジレントアレイ)

ローカルバックグラウンドの半径の自動計算 (True / False で選択)

- **True** の場合は**Self**径を自動計算(デフォルト)
- **False** の場合は値を入力

False
100



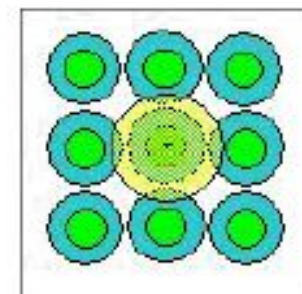


Find and Measure Spots

ローカルバックグラウンドの半径の決定

最小のローカルバックグラウンド半径値 (Self径、デフォルト)

$$0.6 \times \text{Scan Resolution} \times \left[\text{Max} \left(\text{Interspot spacing}_x, \text{Interspot spacing}_y \right) \right]$$

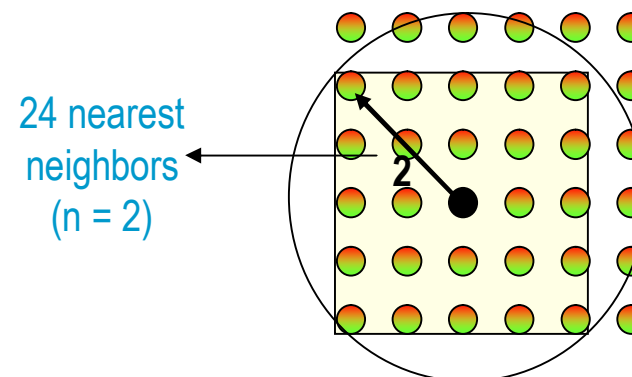


Self

変更可能なローカルバックグラウンド半径値 (最大n = 4)

$$\text{Scan Resolution} \times \left(\sqrt{(n \cdot 6 \times \text{Interspot spacing}_x)^2 + (n \cdot 6 \times \text{Interspot spacing}_y)^2} \right)$$

- スポットからどれぐらいの範囲を直径とするか **n** で決定
- ユーザー設定可能な **n** : 1~4
 - **n = 1** であれば、**8**のスポットを含む近傍エリアから半径を計算
 - **n = 2** であれば、**24**のスポットを含む近傍エリアから半径を計算
 - **n = 3** であれば、**48**のスポットを含む近傍エリアから半径を計算
 - **n = 4** であれば、**80**のスポットを含む近傍エリアから半径を計算





Find and Measure Spots

異常ピクセルの排除法の設定

Protocol Properties

Protocol Steps

Place Grid

Find and Measure Spots

Flag Outliers

Correct Bkgd and Signal Biases

Correct Dye Biases

Compute Ratios and Errors

Calculate Metrics

Generate Results

General

Settings

Use The Nominal Diameter From GridTemplate	True
Spot Deviation Limit	1.50
Pixel Outlier Rejection Method	Inter Quartile Region
Calculation of Spot Statistics Method	Use Cookie
Cookie Percentage	0.650
Exclusion Zone Percentage	1.200
Auto Estimate the Local Radius	True

異常ピクセルの排除法

Standard Deviation
Inter Quartile Region
none

ピクセル排除基準の設定

デフォルト: IQR法
(RejectIQRFeat/BG=1.42)

IQR法

Pixel Outlier Rejection Method	Inter Quartile Region
RejectIQRFeat	1.42
RejectIQRBG	1.42

SD法

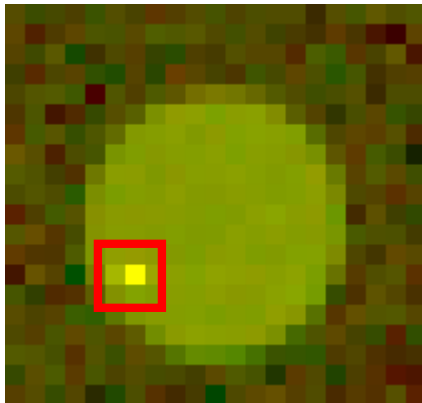
Pixel Outlier Rejection Method	Standard Deviation
Feature SD	2.0
Background SD	2.0



異常ピクセル排除とフラグの違い(概念)

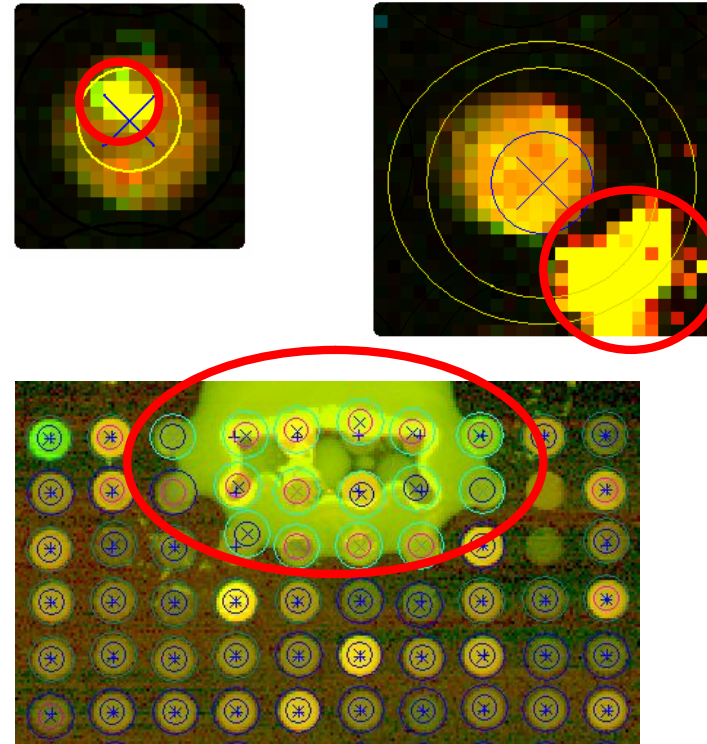
Find and Measure Spots

異常ピクセル排除



Flag Outliers (後述)

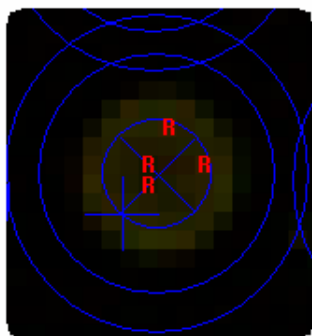
フィーチャとローカルバックグラウンドのフラグ



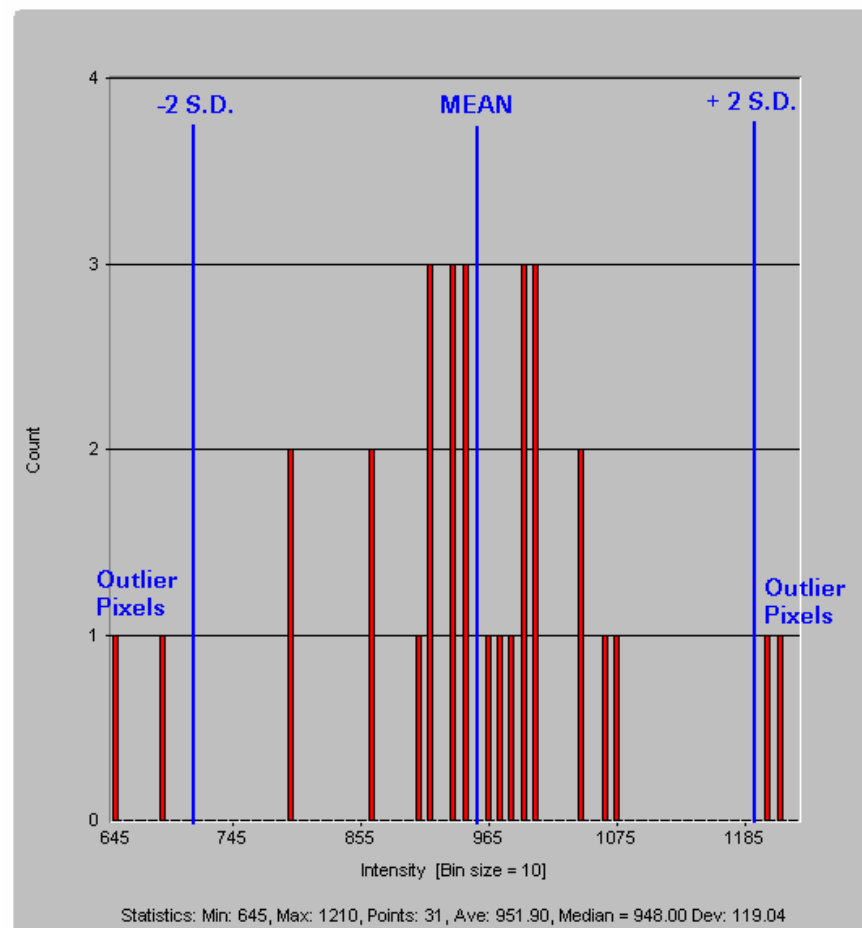


Find and Measure Spots 異常ピクセルの排除

標準偏差 (SD) 法



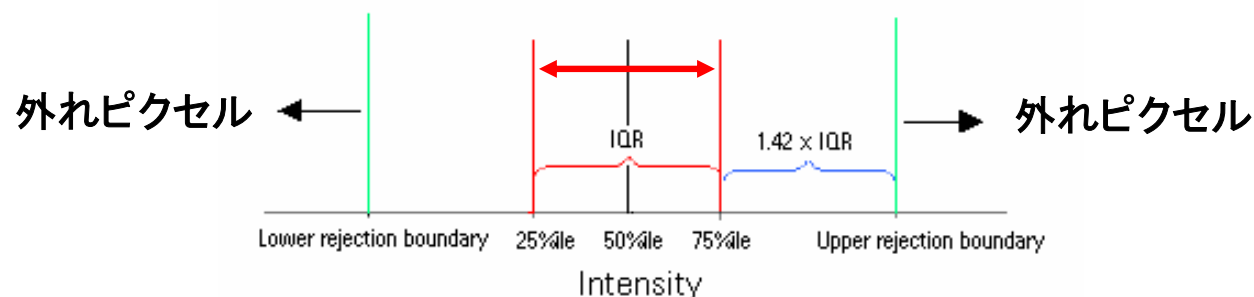
- フィーチャとローカルバックグラウンド領域内の異常ピクセルを排除
- 両チャンネル(赤、緑)で確認
- 標準偏差 $SD \pm 2$ の範囲であれば、
~ 95%の分布を含む
→ $SD \pm 2$ 範囲外のピクセルは異常ピクセルとして排除
- シグナル強度は、 $SD \pm 2$ 以内のピクセルを平均して計算





四分位間範囲 (IQR) 法

四分位間範囲 (IQR) : シグナル強度分布の25-75%間の範囲



- IQR x 1.42 の範囲外を異常ピクセルとして排除
- 正規分布の場合、99%が範囲内におさまる

$$Cutoff_{PixelOutlier} = 1.42 \times IQR$$

$$\text{where } IQR = Intensity_{75th} - Intensity_{25th}$$

- 異常ピクセル排除方法

$$Pixel\ Intensity > RB_{Upper}$$

$$Pixel\ Intensity < RB_{Lower}$$

$$RB_{Lower} = I_{25\ percentile} - Cutoff_{PixelOutlier}$$

$$RB_{Upper} = I_{75\ percentile} + Cutoff_{PixelOutlier}$$



XDR (Extended Dynamic Range)スキャン

①スキャナー



1枚のスライドガラスから、蛍光シグナルの増幅係数の異なる
2枚のイメージが得られる

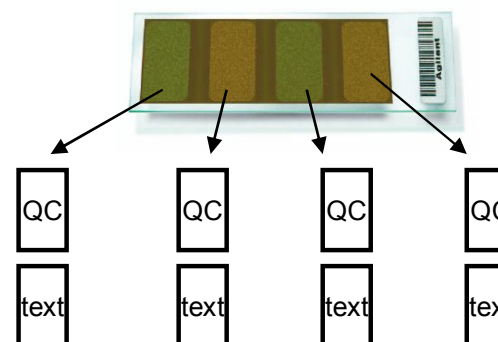


②数値化ソフトウェア



FE9.1以降は、この2つのイメージを
自動的に統合し、1つの結果を出力する

PMT:100 US4510XXXX_2514850XXXX_S01_**H**.tif PMT:10 US4510XXXX_2514850XXXX_S01_**L**.tif



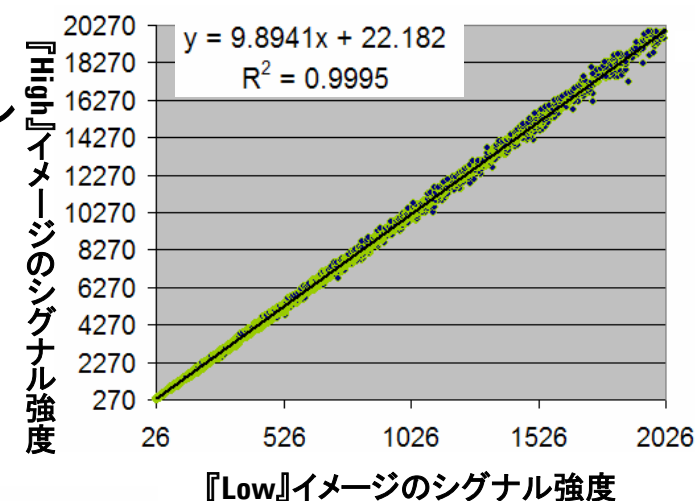
XDRスキャンイメージの統合

Step 1. 2枚のイメージについて、**それぞれNetシグナル強度を抽出**

Netシグナル=ピクセルの平均シグナル強度－スキャナーオフセット(=アレイ上の最低シグナル強度)

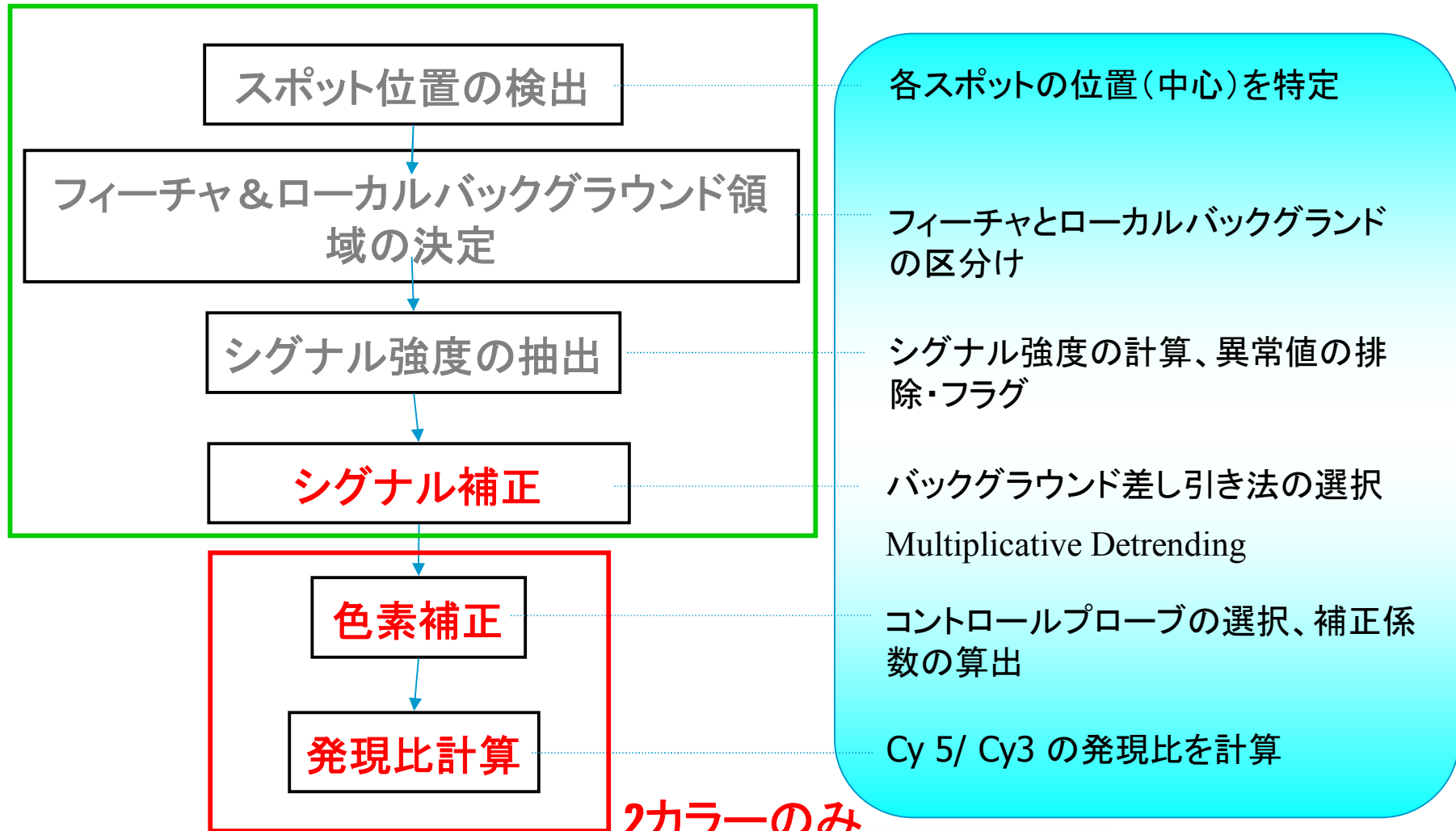
Step 2. 『Low』イメージのNetシグナル強度を
スケールアップするため の、係数・切片を計算
「High」イメージにおいて、**Netシグナル強度が
300から20000の範囲にあるフィーチャーが
算出に用いられる**

Step 3. 2枚の結果を統合して1つのシグナル値を出力
**Highイメージの20000以上のシグナル
強度を持つフィーチャーにおいて、
Netシグナル値が『Low』イメージ
由来のスケールアップ後の値に
置き換えられる**



数値化(スポット解析)の流れ

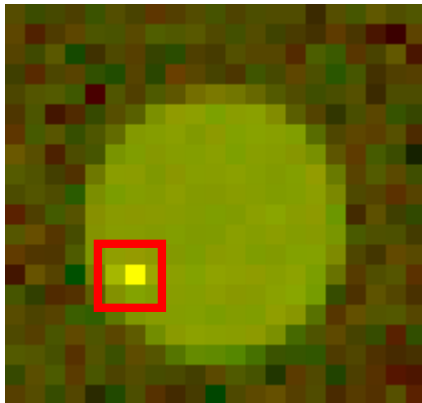
1、2カラー共通



異常ピクセル排除とフラグの違い(概念)

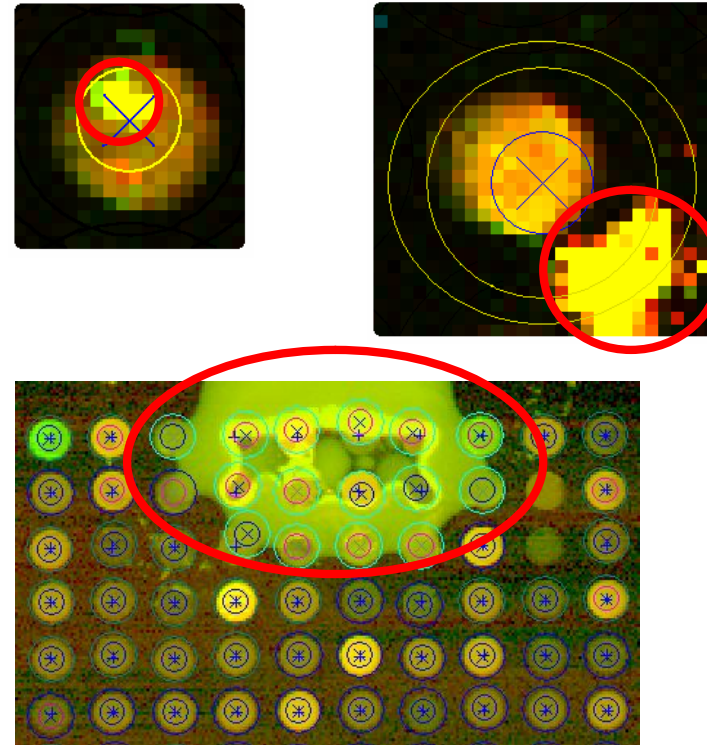
Find and Measure Spots

異常ピクセル排除



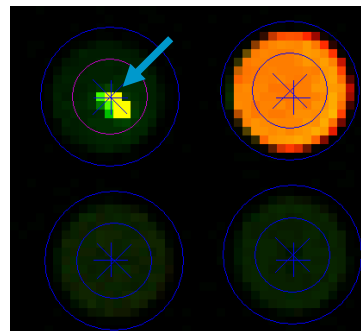
Flag Outliers

フィーチャとローカルバックグラウンドのフラグ





Flag Outliers OutLier(フラグ)の種類



Non-Uniformity

Feature

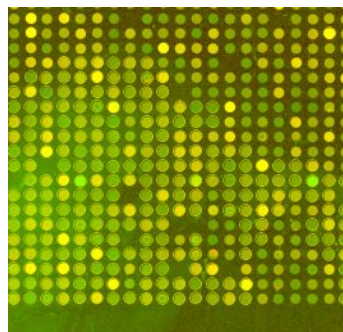
Green

Red

LocalBackGround

Green

Red



Population OL

Feature

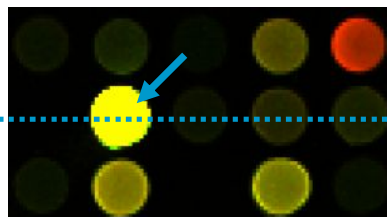
Green

Red

LocalBackGround

Green

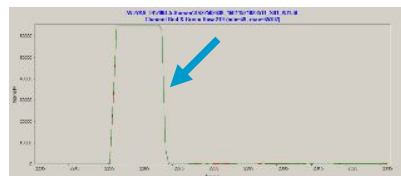
Red



Saturation

Green

Red



※Feature内ピクセルの50%以上saturationした場合





Flag Outliers

アルゴリズム

フラグ自動認識プログラム

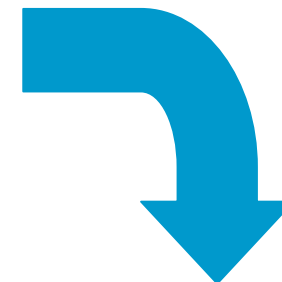
Non-Uniform Outliers

- ・各フィーチャ、ローカルBGについて構成ピクセルの均一性をチェック
- ・アレイ製造、ラベリング、スキャナノイズ等を考慮した多項式分散モデルにより **OutLier** を判定

Population Outliers

- ・繰り返しスポットされたフィーチャ及びローカルBGについて、集団としての均一性をチェック
- ・初期設定の集団の数は **10**
- ・**IQR**または**Q-test** により **Outlier** を判定

- Place Grid
- Find Spots
- Flag Outliers
- Compute Bkgd, Bias and Error
- Calculate Metrics
- Generate Results



General	
Settings	
Compute Population Outliers	True
Minimum Population	10
IQRatio	1.42
Compute NonUniform Outliers	True
Feature	
(%CV)^2 Term	0.00810
Poissonian Noise Term	320
Background Term	600
Background	
(%CV)^2 Term	0.02250
Poissonian Noise Term	320
Background Term	600





Flag Outliers NonUniformity Outlier

Expected Variance(予測値)

$$\sigma_E^2 = Ax^2 + Bx + C$$

$$A = CV^2 = \left(\frac{PixSDev}{MeanSignal - MinSig_{Array}} \right)^2$$

x : (フィーチャまたはバックグラウンドのMean Signal) –
(フィーチャまたはバックグラウンドのMin Signal)

A (Gaussian) – ラベリング及びフィーチャ合成反応から見積もられた分散

B (Poisson) – スキャン測定またはカウントエラーから見積もられた分散

C (Constant) – スキャナの電氣的ノイズ及びスライドガラスのバックグラウンドノイズから見積もられた分散

Measured Variance(実測値)

$$\sigma_M^2 = \frac{1}{n-1} \times \sum_{i=0}^{n-1} (X_i - \bar{X})^2$$

n = フィーチャ または バックグラウンド のInlierピクセルの数

X = フィーチャ または バックグラウンド のピクセルのシグナル強度

$\bar{X}$ = フィーチャ または バックグラウンド のシグナルの平均値

NonUniformity Outlier の判定基準

$$\sigma_M^2 > (\sigma_E^2 + CI) \quad \text{実測値分散が予測値分散よりも大きい場合にフラグ}$$

CI はカイ2乗分布から計算された信頼区間



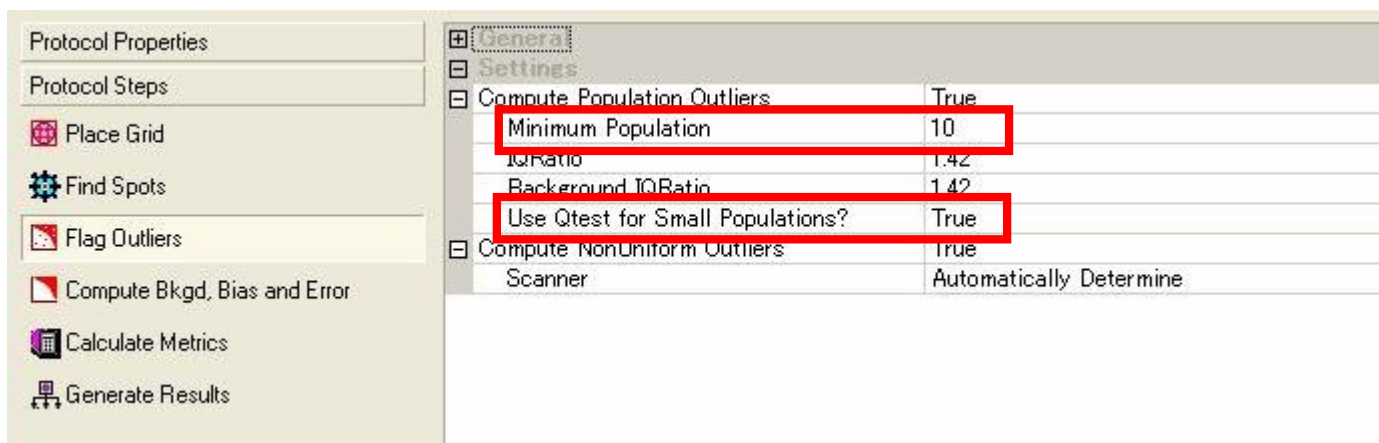


Flag Outliers

Population Outlier

PopulationOutlierの適用対象

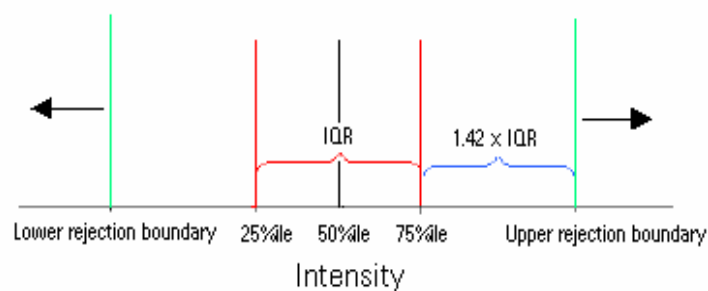
- **1x22K**および**1x44K**、あるいは**Feature Extraction9.1**の場合
→ 繰り返しスポット数: 10スポット以上、**Outlierの認識法: IQR**
- **4x44K**、**8x15K**および**1x244K (Feature Extraction9.5)**の場合
 - 繰り返しスポット数: 11スポット以上、**Outlierの認識法: IQR**
 - 繰り返しスポット数: 3~10スポット、**Outlierの認識法: Q-test**





Flag Outliers

Population Outlier



$$Cutoff_{PixelOutlier} = 1.42 \times IQR$$

where $IQR = Intensity_{75th} - Intensity_{25th}$

1,42*IQRを使用した場合、99 %以上の
分布が排除限界の内側に含まれる。

Population outlier の判定基準

$$MeanSignal > RB_{Upper}$$

$$MeanSignal < RB_{Lower}$$

$$RB_{Lower} = I_{25\text{percentile}} - Cutoff_{PopOutlier}$$

$$RB_{Upper} = I_{75\text{percentile}} + Cutoff_{PopOutlier}$$





Flag Outliers

Population Outlier

Q-test

$Q_{exp} > Q_{critical}$,then FeatPopOutlier = 1

$$Q_{exp} = |X_i - X_{nearest}| / |X_{max} - X_{min}|$$

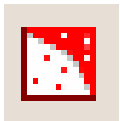
X_i : あるプローブのシグナル強度

$X_{nearest}$: X_i に一番近いシグナル強度

繰り返しスポット数	$Q_{critical}$
3	0.970
4	0.829
5	0.710
6	0.625
7	0.568
8	0.526
9	0.493
10	0.466

95% Confidence level

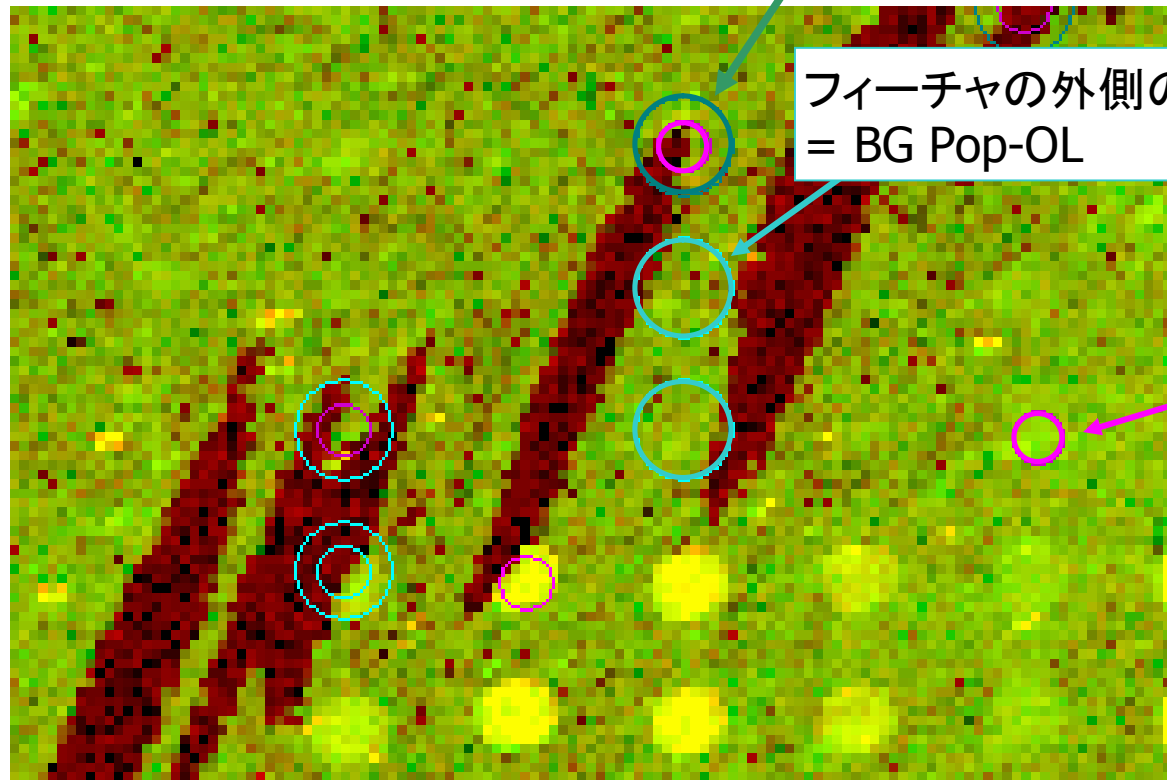




Flag Outliers

画面上での OutLier の確認

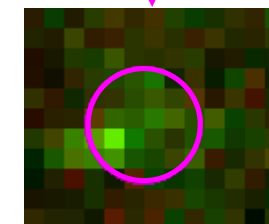
ログスケール



フィーチャの外側の緑のリング
= BG non-uniformity

フィーチャの外側の青緑のリング
= BG Pop-OL

フィーチャの内側の
ピンクのリング
= Feature non-uniformity



リニアスケール



Compute Bkgd, Biases and Error

アルゴリズムの流れ



1. 選択された **Background Subtraction** を実施
2. **Spatial detrend** を実施 (**True**の場合)
3. **Error値**の計算
(伝播エラーモデルとユニバーサルエラーモデルのいずれかを使用)
4. **有意差検定**を実施
(**Significance test**と**Well above BG**)
5. **Surrogate値**の計算
6. **Multiplicative Detrending**

オプション

オプション

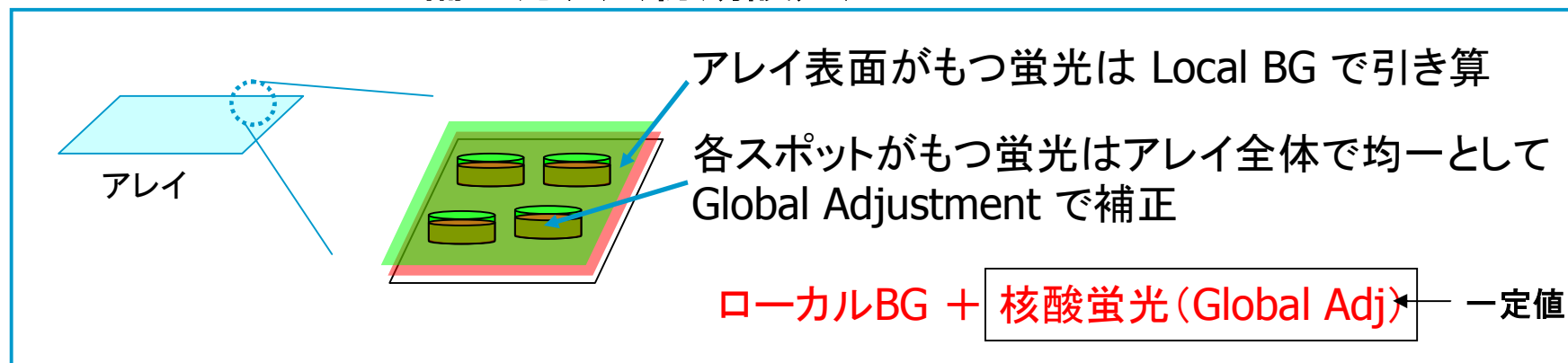




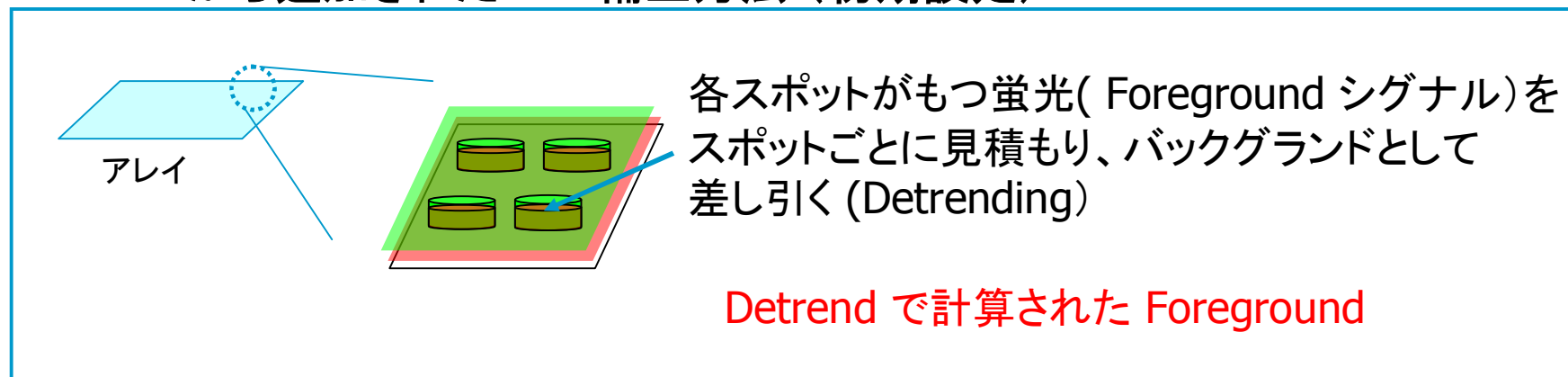
Compute Bkgd, Biases and Error

バックグラウンド補正について

FE 7.1 までの BG 補正方法（初期設定）



FE 7.5 から追加された BG 補正方法（初期設定）





Compute Bkgd, Biases and Error

1. Background Subtraction

No Background Subtraction (初期設定)

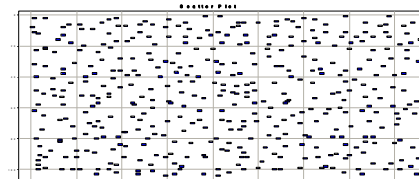
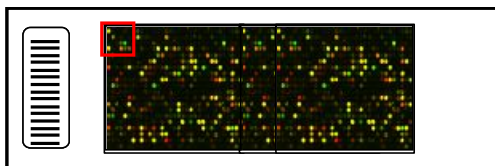
- フィーチャのシグナルから ローカル **Background** のシグナルを 差し引かない
- フィーチャの生シグナル値 (**MeanSignal**) を次のステップの **Spatial Detrend** に使用 (**True**の場合)
- **No Background Subtraction** を選択した場合、**Global Adjustment** は行わない (初期設定)





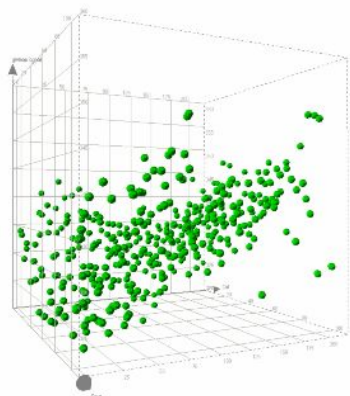
Compute Bkgd, Biases and Error

2. Spatial Detrend の流れ

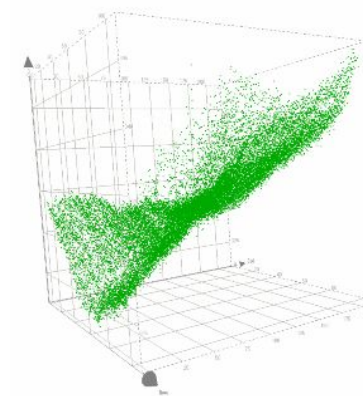


1. アレイ上でウィンドウをスキャンさせ、ウィンドウ内で、一定の低いシグナル強度範囲にあるフィーチャを検出

2. 検出されたフィーチャから **Foreground** シグナルの見積もりに使用するフィーチャを決定
(CGH・ChIP on chipではネガコンを使用)



3. 得られた **Foreground** シグナルの分布を **2D-LOESS** アルゴリズムを用いてフィッティング



4. 全フィーチャに対して各々の **Foreground** シグナルを算出し、**Mean**シグナルから差し引く





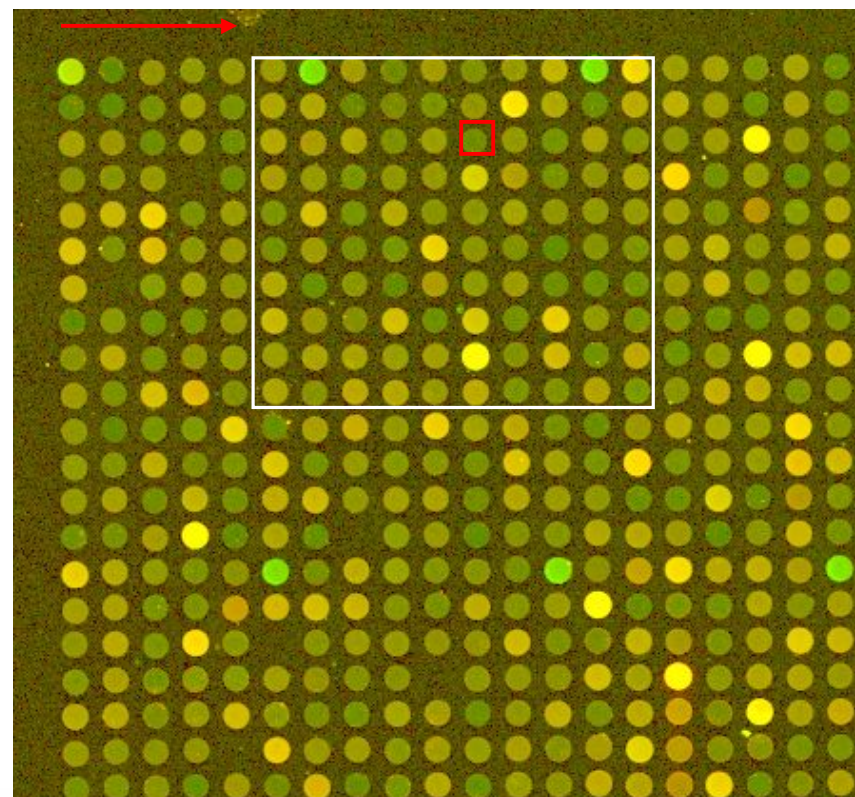
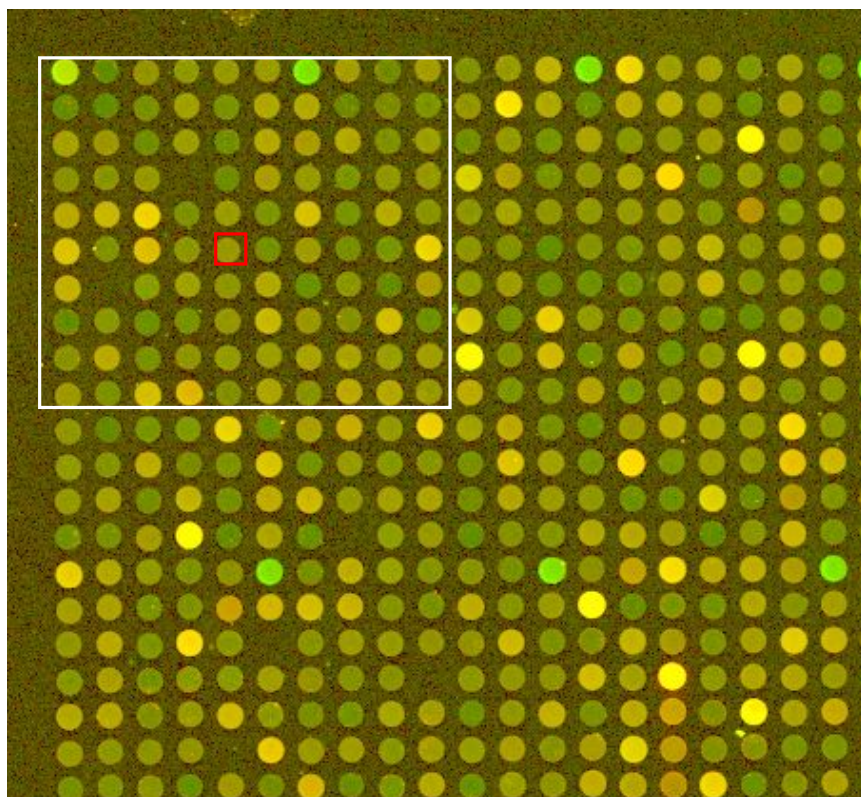
Compute Bkgd, Biases and Error

2. Spatial Detrending

ウィンドウの移動

各ウィンドウごとに一定のシグナル範囲にあるフィーチャを検出

Window = 10, Increment = 5



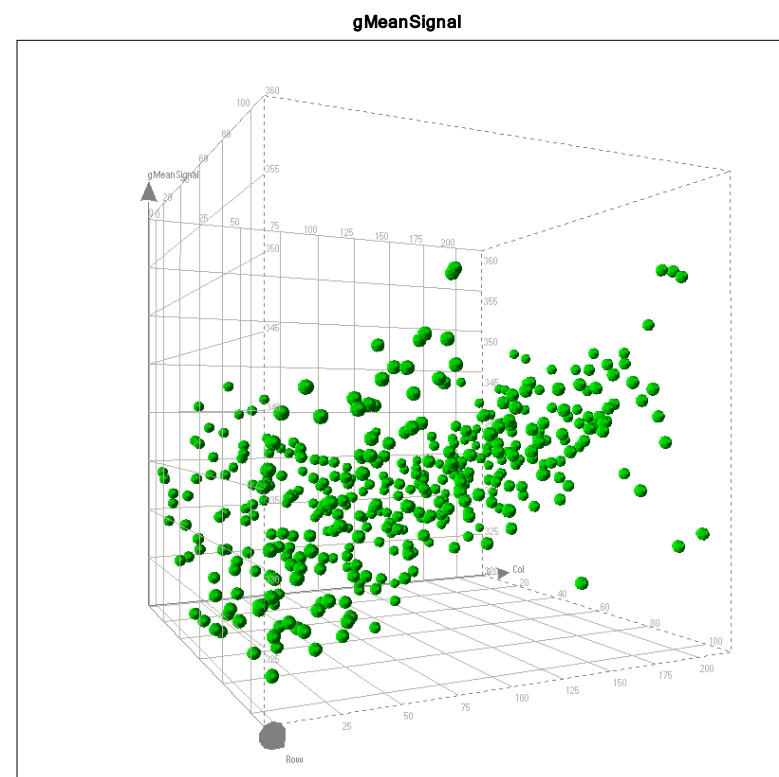
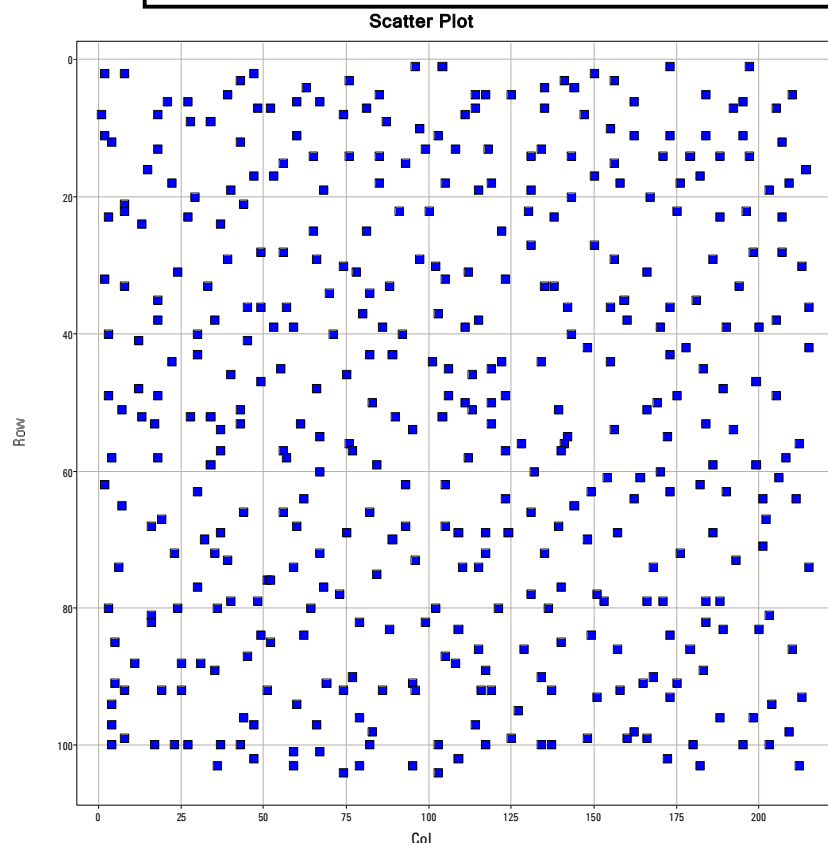


Compute Bkgd, Biases and Error

2. Spatial Detrending

ウィンドウ移動により選択されたフィーチャセットと **Foregroundシグナルの強度分布(拡大図)**

Features from Low Pass Filter – Raw Green Intensities





Compute Bkgd, Biases and Error

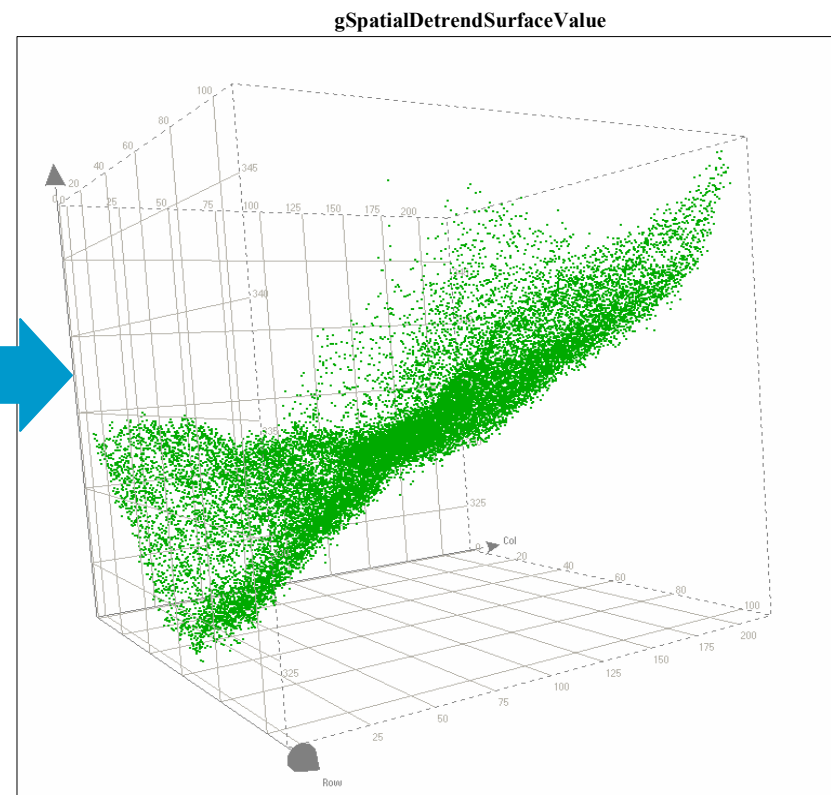
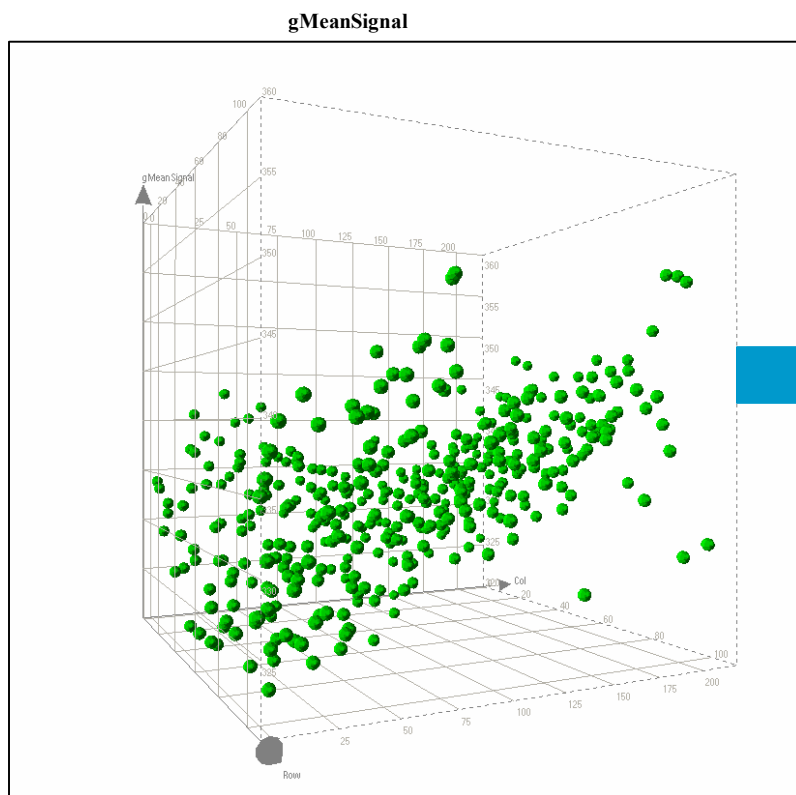
2. Spatial Detrending

2D Loess フィット

全てのフィーチャに対する **Foreground** シグナル値の算出（拡大図）

Filtered set

2D Loess Fit

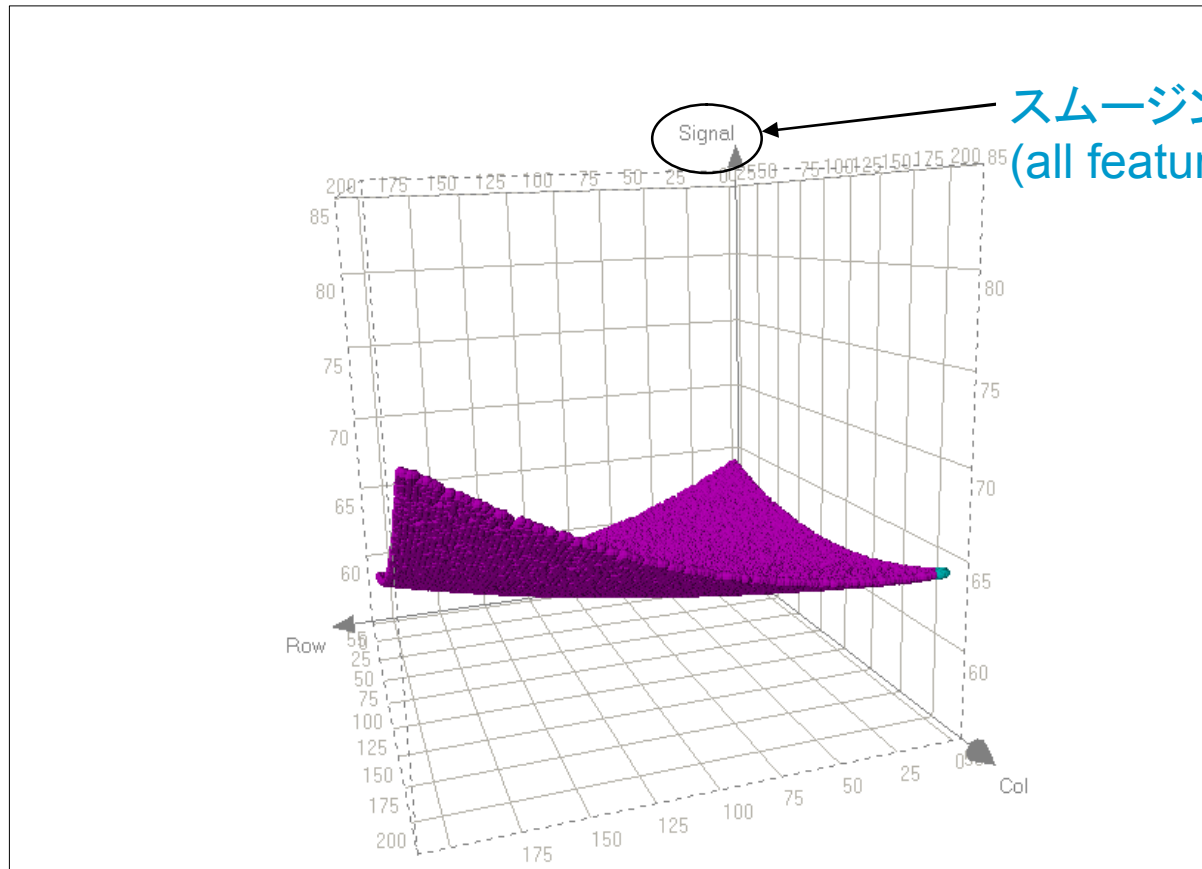




Compute Bkgd, Biases and Error

2. AgilentアレイのSpatial Detrending

Fitted Smoothed meanSignal (~60 - ~85 counts)



スムージング後のシグナル強度
(all features)



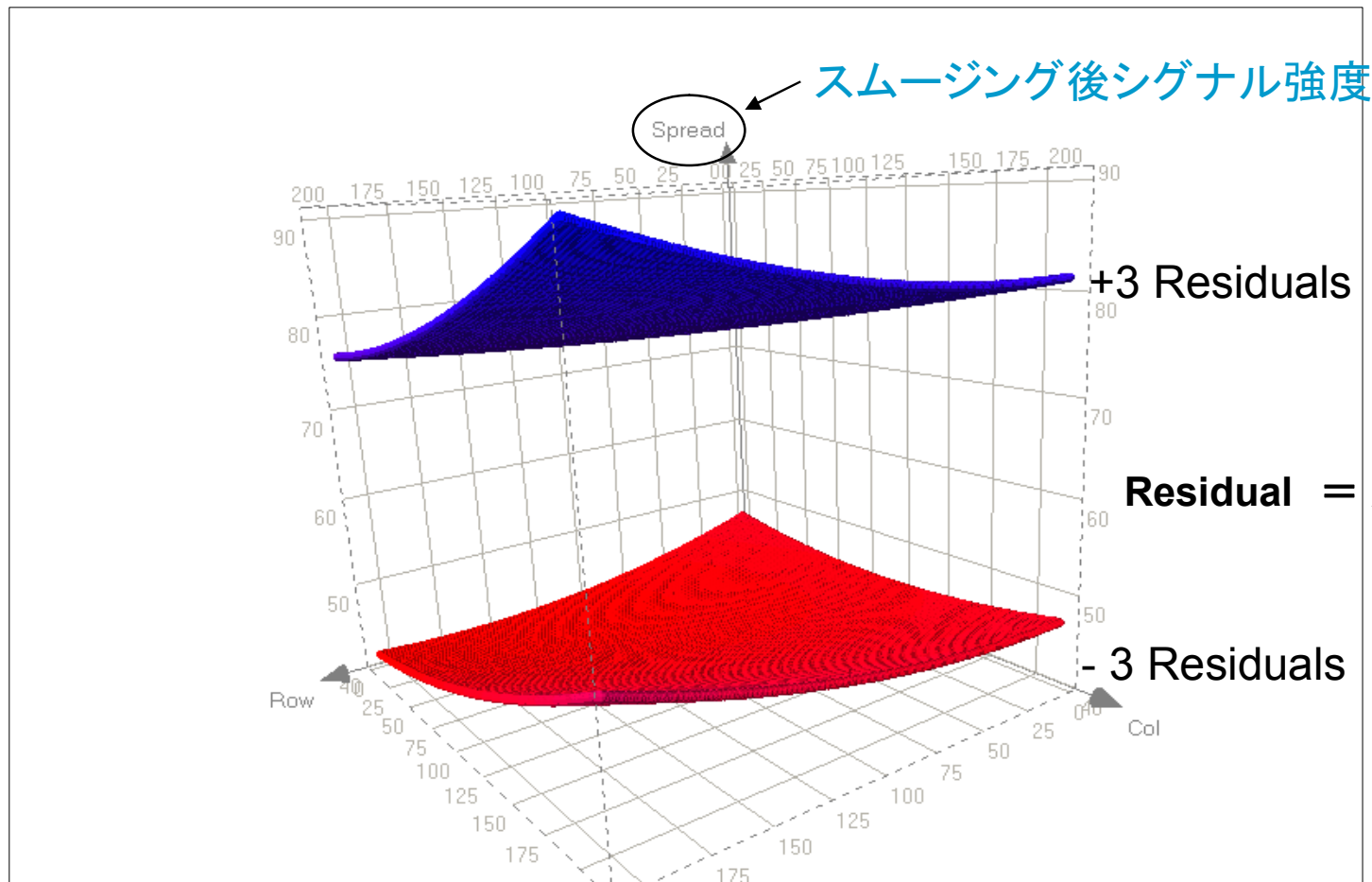


Compute Bkgd, Biases and Error

2. Agilent アレイのSpatial Detrending

残差分析を行い、外れ値を除く

Spread of Negative Controls (~50 - ~90 counts)



I = Expected Fit Value
O = Observed Fit Value

$$\text{Residual} = \sqrt{\frac{\sum_{i=1}^N (I_i - O_i)^2}{N}}$$

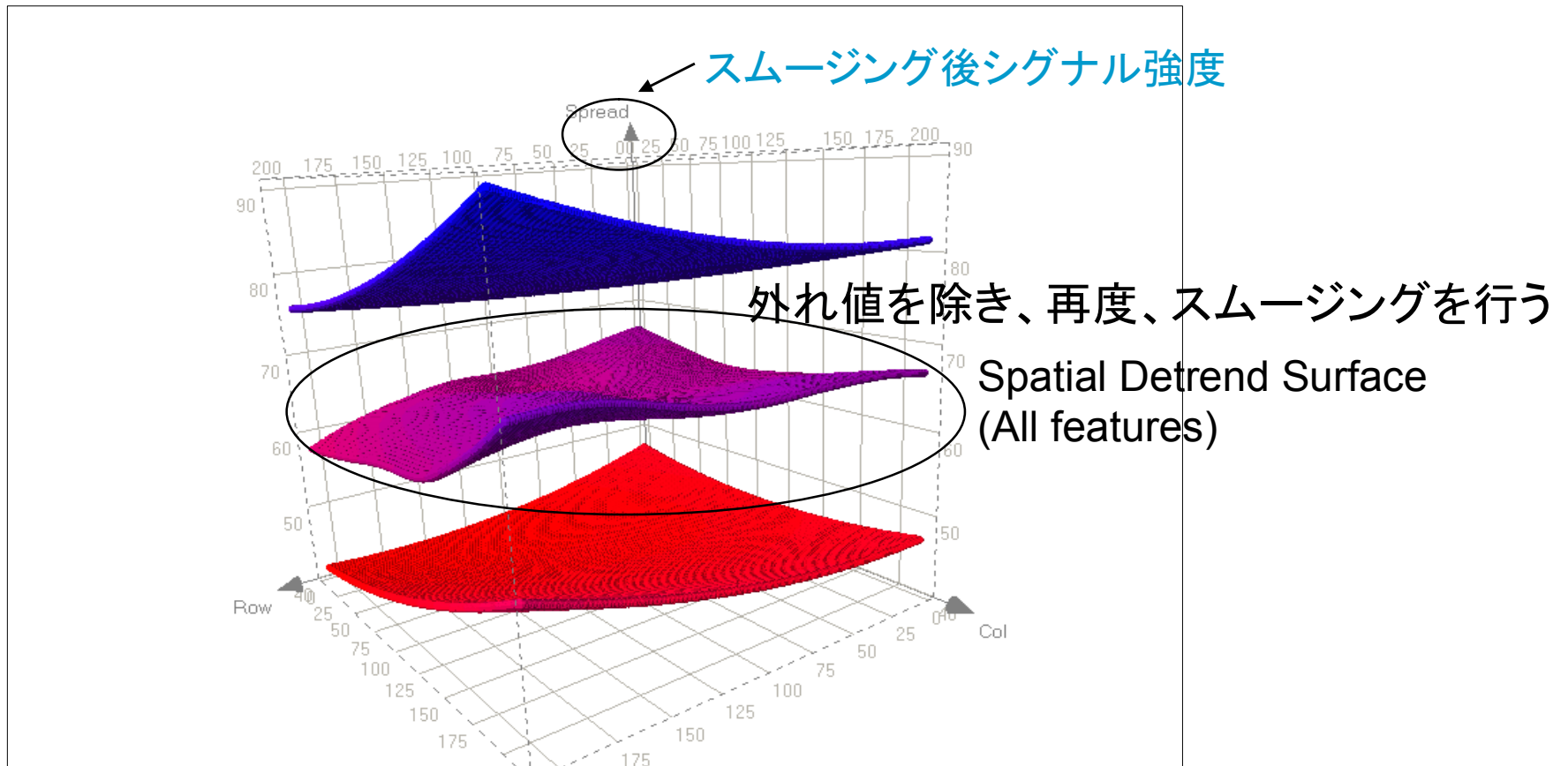




Compute Bkgd, Biases and Error

2. Agilent アレイのSpatial Detrending

Spatial (Additive) Detrend Surface





Compute Bkgd, Biases and Error

2.Spatial Detrendを使用した時の出力項目

詳細はテキストアウトプットを参照

FEATURES table (2 項目):

[r,g]SpatialDetrendIsInFilteredSet – 0 または 1

- **1 = foreground** シグナルの算出のために選択されたフィーチャ
- フラグ(**FeatNonUnifOLs, BGNonUnifOLs**)・サチュレーション・**Pro25G** コントロールフィーチャは **foreground** シグナル算出に使用しない

[r,g]SpatialDetrendSurfaceValue

- 各フィーチャの **foreground** シグナルの値
(**LOESS**を用いた **surface fit** によって決定)





Compute Bkgd, Biases and Error

2.BGSubSignal 計算方法まとめ

$$\text{BGSubSignal} = \text{MeanSignal} - \text{BGUsed}$$

Table 29 Values for BGSubSignal, BGUsed and BGSDUsed for different methods and settings*

Background Subtraction Method	Background Subtraction Variable	Spatial Detrend (SpDe) OFF Global Bkgnd Adjust (GBA) OFF	SpDe ON GBA OFF	SpDe OFF GBA ON	Spatial Detrend ON Global Bkgnd Adjust ON
No background subtract	BGUsed =	BGMeanSignal [†]	SpatialDetrendSurfaceValue	BGAdjust	SpatialDetrendSurfaceValue (SDSV) + BGAdjust
	BGSDUsed =	BGPixSDev	BGPixSDev	BGPixSDev	BGPixSDev
	BGSubSignal =	MeanSignal	MeanSignal - BGUsed	MeanSignal - BGUsed	MeanSignal - BGUsed
Local Background	BGUsed =	BGMeanSignal	BGMeanSignal + SDSV	BGMeanSignal + BGAdjust	BGMeanSignal + SDSV + BGAdjust
	BGSDUsed =	BGPixSDev	BGPixSDev	BGPixSDev	BGPixSDev
	BGSubSignal =	MeanSignal - BGUsed	MeanSignal - BGUsed	MeanSignal - BGUsed	MeanSignal - BGUsed
Global Background method	BGUsed =	GlobalBGInlierAve (GBGIA)	GBGIA + SDSV	GBGIA + BGAdjust	GBGIA + SDSV + BGAdjust
	BGSDUsed =	GlobalBGInlierSDev (GBGISD)	GBGISD	GBGISD	GBGISD
	BGSubSignal =	MeanSignal - BGUsed	MeanSignal - BGUsed	MeanSignal - BGUsed	MeanSignal - BGUsed

* For both the red and green channels

† With No background subtraction as the setting, BGMeanSignal is the value for BGUsed only for the t-test, but no BGUsed is subtracted from the MeanSignal to produce BGSubSignal.





Compute Bkgd, Biases and Error

3. エラー値の計算

エラーモデル

Log Ratio のエラーを見積もるための3つのエラーモデル

- ピクセルレベルの統計に基づく伝播エラーモデル
(Agilent's propagated error method based on pixel-level statistics)
- ロゼッタ社のユニバーサルエラーモデル
(Rosetta's Universal Error Model: UEM)
- 伝播エラーモデルと**UEM**から導かれるエラーのうち、
より大きな値を採用するモデル(デフォルト)

採用するエラーモデルに応じて、**Log Ratio = 0** の可能性(**P**値)を計算

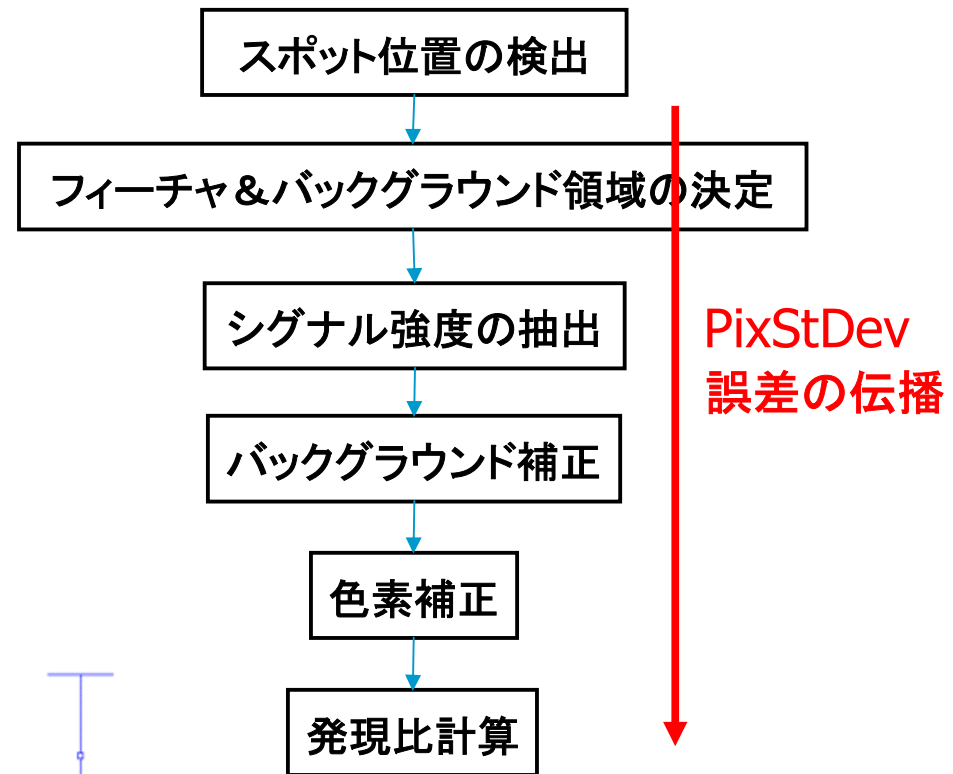
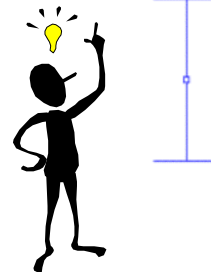


Compute Bkgd, Biases and Error

3. エラー値の計算

伝播エラーモデル

- 各アルゴリズム (例: シグナル抽出、バックグラウンド補正等) のステップで生じるピクセルレベルの誤差の伝播に基づき **Log Ratio** の誤差 (エラー) を算出
- シグナルの低い領域において有効
- シグナルが中程度およびより高い領域において過小評価





Compute Bkgd, Biases and Error

3. エラー値の計算

ユニバーサルエラーモデル

Additive および **multiplicative** エラー項を用いて赤および緑チャンネルで期待されるエラーを算出

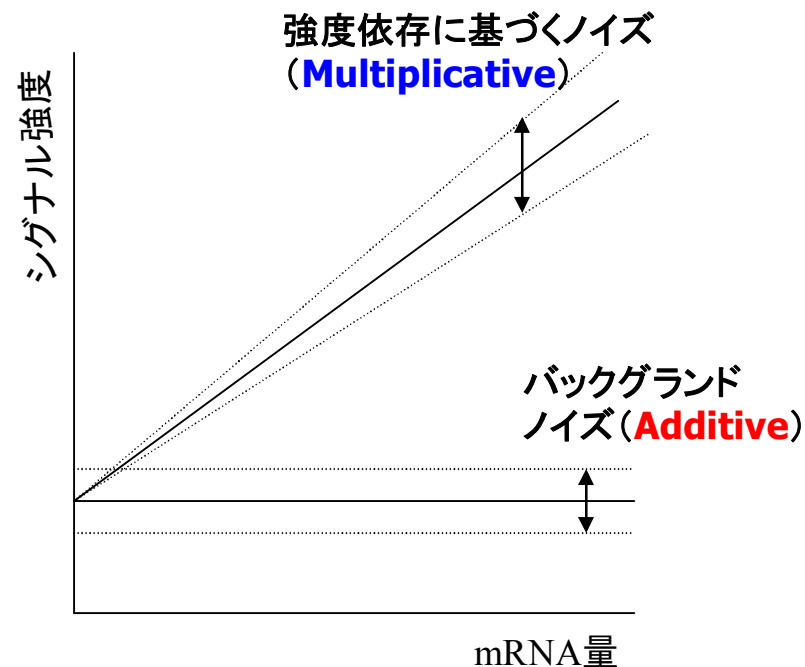
- **Additive**

一定量のノイズエラー
(低シグナル領域で優先)

- **Multiplicative**

シグナル強度依存の変動に基づくエラー
(高シグナル領域で優先)

シグナルが中程度およびより高い領域において有効ノイズのあるフィーチャ、特にシグナルの低い領域において過小評価



伝播エラーモデルとユニバーサルエラーモデルのうち
エラーの大きい方を使用 (**Conservative: 推奨**)

両方のエラーモデルを評価し、より大きな(より保守的な)P値をレポート





Compute Bkgd, Biases and Error

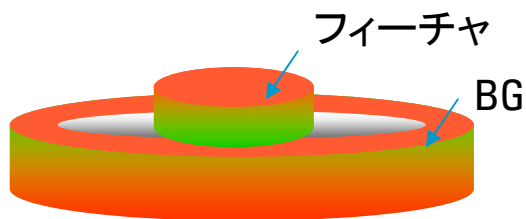
4.フィーチャシグナルの有意差検定

Feature Significance Test

BGシグナルもしくはエラーの値に対し、フィーチャシグナルがsignificantであるかどうかを判定:

2-sided Student's t-test

$$pValue_{Calculated} < pValue_{Max}$$



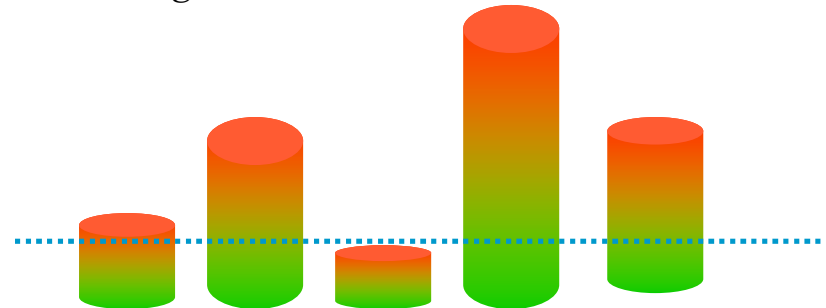
計算されたp-valueがユーザーが設定したp-value_{max}より小さい場合は、FE出力のIsPositiveAndSignif の項目で **1** を出力

※Agilent高密度フォーマットアレイでは、t-testにローカルBGの分散ではなく**Error値**を使用

Well Above Background Test

BGSubシグナルが、下式で計算されたバックグラウンド値よりも “well above” であるかどうかを判定:

$$BGSubSignal > WellAboveSDMulti \times BGSDUsed$$



IsPositiveAndSignif testとWellAboveの両方をパスしたフィーチャに対して、FE出力のIsWellAboveBG column の項目で1を出力

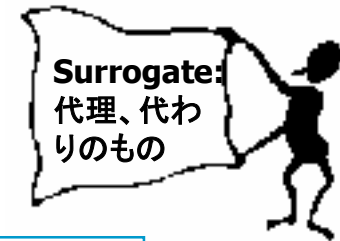




Compute Bkgd, Biases and Error

5. Surrogate値の計算

Surrogateとは...



- 発現量が少なく、生物学的意味のないシグナル強度を示している値
- 小さすぎて、その後の計算に支障をきたす値

具体的には、これまでの過程を経たシグナルが、下記のいずれかに該当する場合...

- **BGSubSignal** が **IsPosAndSignif** の判定ではねられる(**Fail**する)場合
- **BGSubSignal** が バックグラウンドの標準偏差 (i.e. **BGSDUsed**) よりも小さい場合



サロゲート値(バックグラウンドレベルの値)に置き換える

サロゲート値: バックグラウンドシグナル強度の **1SD** (**x DyeNormFactor** (2色のみ))

(前ステップの有意差検定に使用した方法によって、値は異なる)

- ローカルバックグラウンド法 → ピクセルレベルのローカルバックグラウンドの **SD**
- グローバルバックグラウンド法 → アレイ上の集団(**population**)レベルの **SD**
- **Error Model**法 → (**Additive Error**値 / **2.6**) × **P**値





Compute Bkgd, Biases and Error

5. Surrogate値の計算

サロゲート値が適用された場合、FEアウトプットテキストファイルの『**Surrogate Used**』にサロゲート値が表示される

(適用されなければ0)

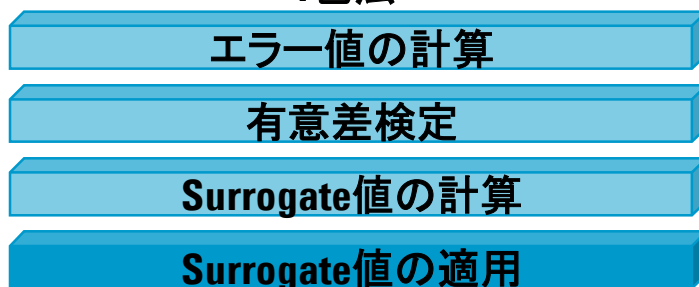
P	U	R
:	float	float
RowSt	Grid_ColSpacing	Grid_Offse
3.3235	63.5	0
:	integer	
balBGI	gGlobalBGInlierNu	
11522	4	
:	float	boolean
itionY	gSurrogateUsed	gIsFound
31.21	0	1
19	3.21907	0
31.06	0	1
19	3.20219	1
31.06	3.1944	1

FE output で
確認可能



※2色法では、色素補正後に
Surrogate値が適用される

1色法



2色法

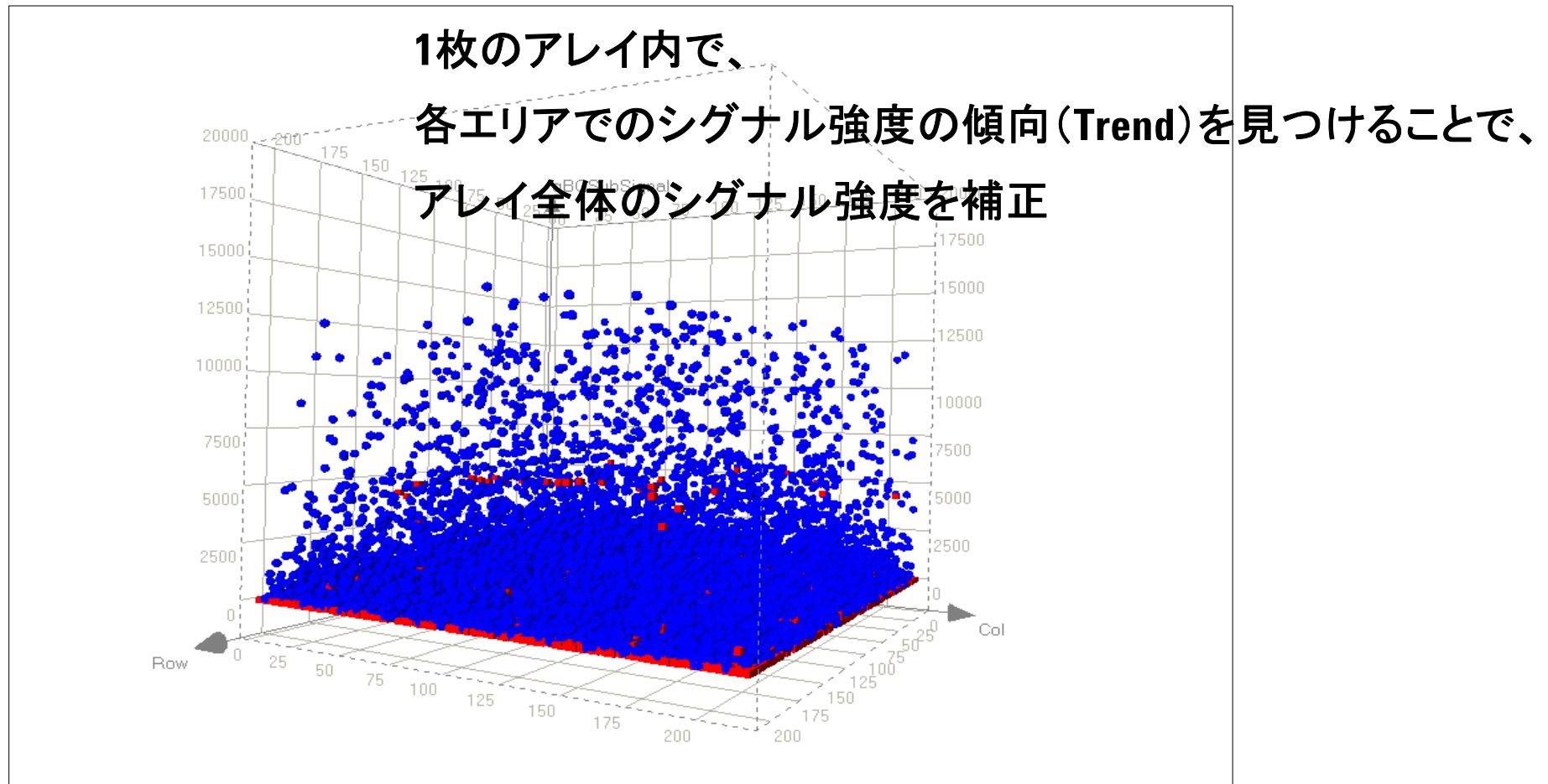




Compute Bkgd, Biases and Error

6. Multiplicative Detrending – Finding the Trend

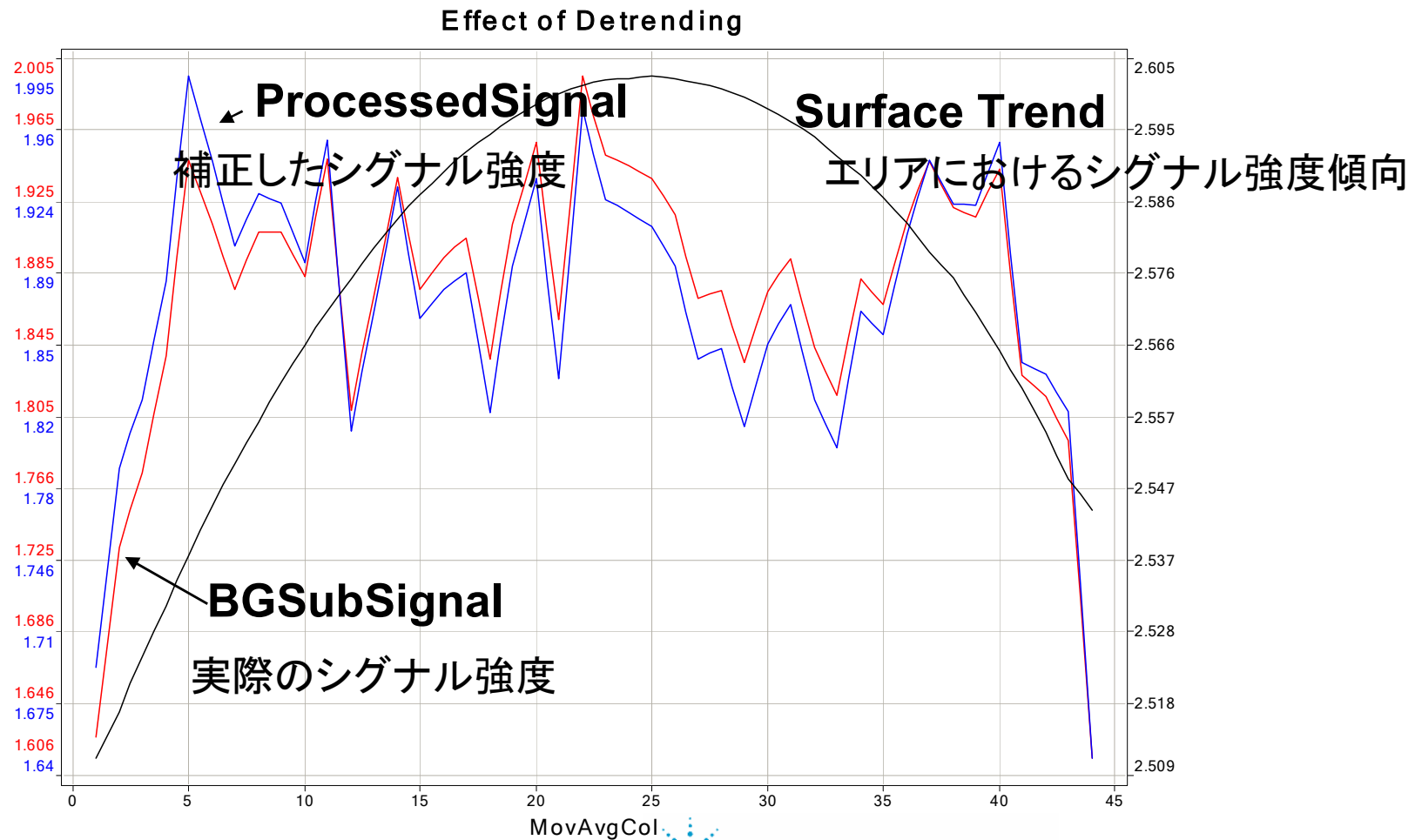
Look For Trends by Fitting The BGSubSignal





Compute Bkgd, Biases and Error

6. Multiplicative Detrending





Compute Bkgd, Biases and Error

6. Multiplicative Detrending - WARNING

レプリケート(繰り返し)プローブのCV値を指標に
Multiplicative DetrendingのOn/Offを自動で決定
→Offの場合、**WARNING**が出現

Fri Oct 28 11:22:06 2005

INFO: Grid in use: 012790_D_20041018.

INFO: Protocol in use: GE1_22k_1105. |

INFO: There are 4380 (Green) negative features (Non-Controls).

WARNING: Multiplicative detrending effect inconclusive (CVs increasing): detrending removed.

INFO: 129 (Green) saturated features

INFO: 161 (Green) feature non-uniformity outliers

INFO: 59 (Green) feature population outliers

INFO: 3 (Green) background non-uniformity outliers

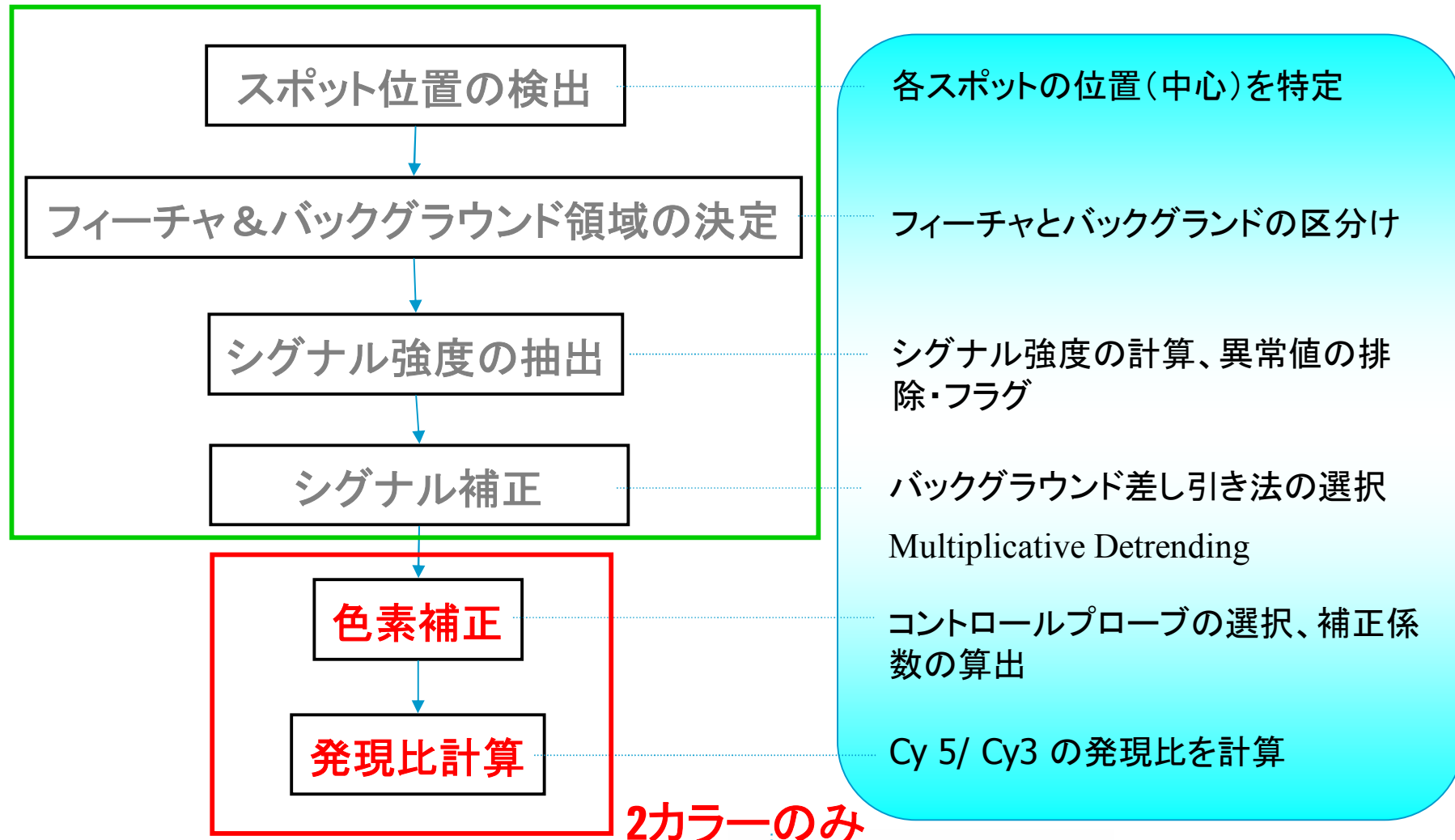
INFO: 1375 (Green) background population outliers

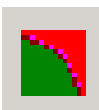
- MD前のCV値が、MD後のCV値よりも大 → MDがOn
- MD前のCV値が、MD後のCV値よりも小 → MDがOff



数値化(スポット解析)の流れ

1、2カラー共通





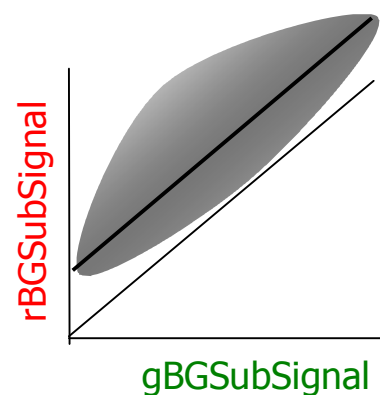
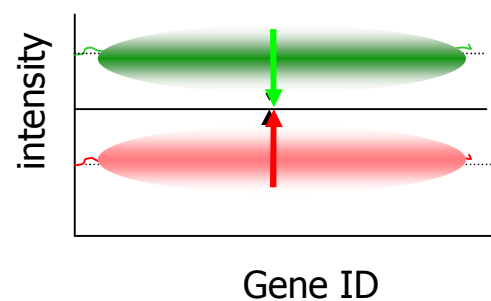
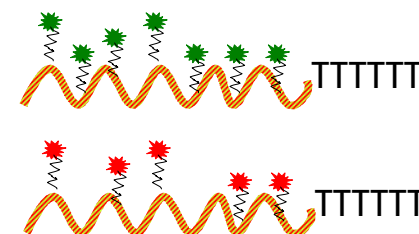
Correct Dye Biases

色素バイアスの概念



色素バイアスの要因(例)

- 2サンプル間の**RNA量**が不揃い
- 2つの色素のラベリング効率の相違
- スキャン時のレーザーパワーの違い





色素バイアス補正のアルゴリズム:

Dye Normalization Probe Selection Method	Use Rank Consistent Probes
Rank Tolerance	0.050
Omit Background Population Outliers	False
Allow Positive and Negative Controls	False
Signal Characteristics	OnlyPositiveAndSignificantSignals
Normalization Correction Method	Linear And Lowess

1. 補正係数の計算に用いるフィーチャの選択

- **Rank Consistency Filter**
- **Use All Probes**
- **Use List of Normalization Genes**
- **Use Rank Consistent List of Normalization Genes**

2. 補正係数(normalization factor)の計算

- **Linear**
- **Linear&LOWESS**
- **LOWESS**

1. パーセンタイルの閾値

2. このオプションを選択すると、BGPpnOL も考慮される(デフォルトはFalse)。

3. このオプションを選択すると、コントロールプローブも考慮される(デフォルトはFalse)。

4. バックグラウンドと比べて有意または有意でないと判断されたプローブセットを考慮するかどうかを選択可能(デフォルトはPosAndSignifプローブのみを使用)。





Correct Dye Biases

補正係数の計算に用いるフィーチャの選択

Use Rank Consistent Probes (デフォルト)

- 赤チャンネルと緑チャンネル間で “**Central Tendency Line (中心線)**” に沿って位置するフィーチャを使用
 - リアルタイムに ハウスキーピング遺伝子を選択するような作業。

Use All Probes

- コントロールでなく、且つ異常値を持たないフィーチャを使用
 - Non-control → ControlType = 0
 - Non-outlier → 両チャンネルで IsFeatNonUnifOL, IsFeatNonPopnOL 且つ 飽和していない

Use List of Normalization Genes

- ハウスキーピング遺伝子あるいは発現差が無いと考えられる遺伝子を使用(ユーザーが定義:外部ファイルを使用可)

Use Rank Consistent List of Normalization Genes

- 定義する遺伝子リストに対し、Rank Consistency Filterを適用





Correct Dye Biases

Rank Consistency Filter

中心線に沿ったフィーチャの同定

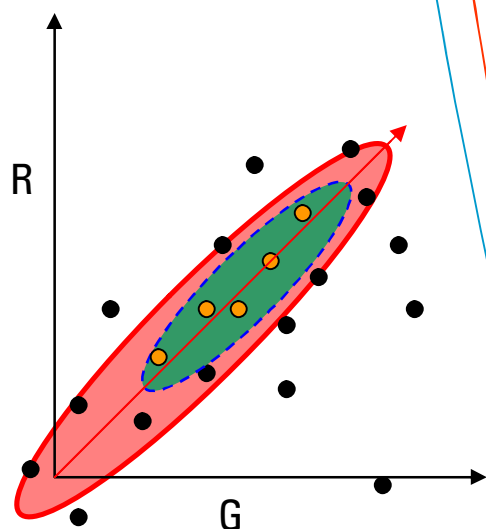
各フィーチャのシグナル強度をチャンネル毎にランク付けから計算。

各フィーチャのチャンネル別ランクに基づく中心傾向 = $|\rho_R - \rho_G| / \Sigma(\text{Features}) \leq \tau$
(τ : パーセンタイル閾値)

両チャンネルのランクをフィーチャ毎に比較し、 τ パーセンタイル以内に入るフィーチャを同定

例: 中心傾向5パーセンタイル($\tau = 0.05$) 以内に入っているフィーチャ

中心傾向15パーセンタイル($\tau = 0.15$) 以内に入っているフィーチャ



フィーチャ番号	赤チャンネル強度 I_R	ランク(赤) ρ_R	緑チャンネル強度 I_G	ランク(緑) ρ_G
1	30	1	560	5
2	170	5	390	4
3	99	4	146	1
4	360	6	452	6
5	43	2	300	3
6	45	3	149	2
7	423	7	700	7

赤と緑で
似通ったラ
ンクである
フィーチャ
は、...



フィーチャ選択のための付加的なフィルタ:

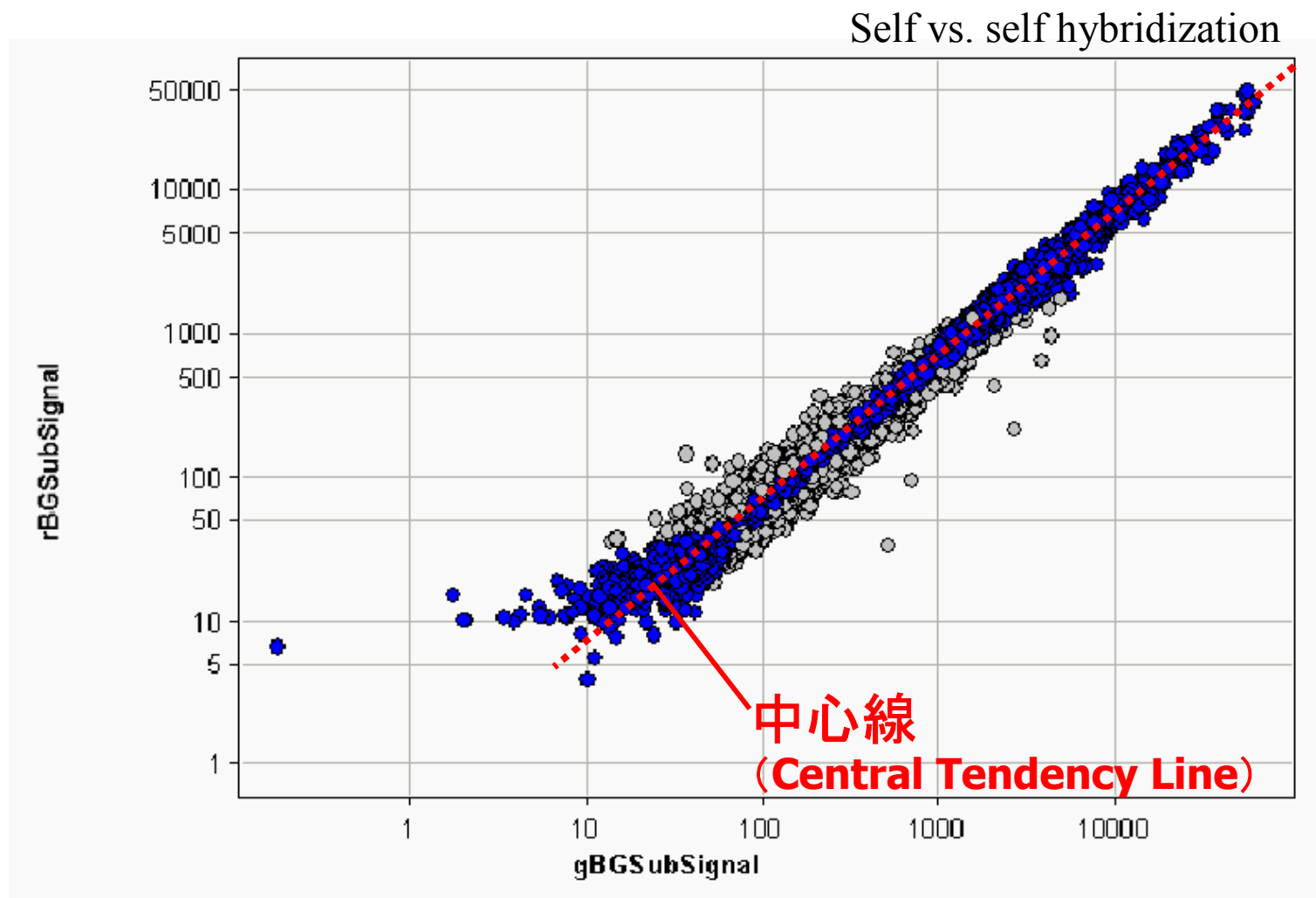
1. Not Saturated
2. Non Population Outlier
3. Non Non-Uniformity Outlier





Correct Dye Biases

Rank Consistency Filter により選択されたフィーチャ(青色)





Correct Dye Biases 補正係数の決定 **Linear**

Linear ノーマライゼーション法

色素バイアスはシグナル強度に関わらず一定と仮定(グローバル)
平均の **log ratio** はゼロ

赤と緑の各チャンネルで補正係数(**LinearDyeNormFactor**)を決定

補正係数の決定の為に選択されたフィーチャの幾何平均が **1,000** となる様に係数を計算

例: 幾何平均が **250** の場合、**LinearDyeNormFactor = 4**

リニア法の問題点 → 色素バイアスがシグナル強度依存的な場合は不適切





Correct Dye Biases 補正係数の決定 **LOWESS**

LOWESS ノーマライゼーション法

LOWESS : locally weighted linear regression

(局所重み付け直線回帰)

色素バイアスはシグナル強度依存的と仮定

赤と緑の各チャンネルで各フィーチャに対し、強度依存的な補正係数
(**LOWESS** または **Linear&LOWESS DyeNormFactor**)を決定

補正係数の決定の為に選択されたフィーチャ(ノーマライゼーションセット)に対し、
局所重み付けの直線回帰によるフィッティングを実施





Correct Dye Biases **DyeNormFactor (DNF) の計算**

Linear 法

$$LinearDyeNormFactor = \frac{1000}{10^{\left(\frac{1}{n} \sum_{i=1}^n \log_{10} BGSubSignal_i\right)}}$$

n はノーマライゼーションセットにおける
フィーチャの数 (i.e. [IsNormalization = 1](#)
のフィーチャ)

FE output で
確認可能



Linear&LOWESS 法: (アジレントアレイのデフォルト)

$$Linear \ \& \ LOWESSDyeNormFactor = \frac{DyeNormSignal}{BGSubSignal \times LinearDyeNormFactor}$$

LOWESS 法

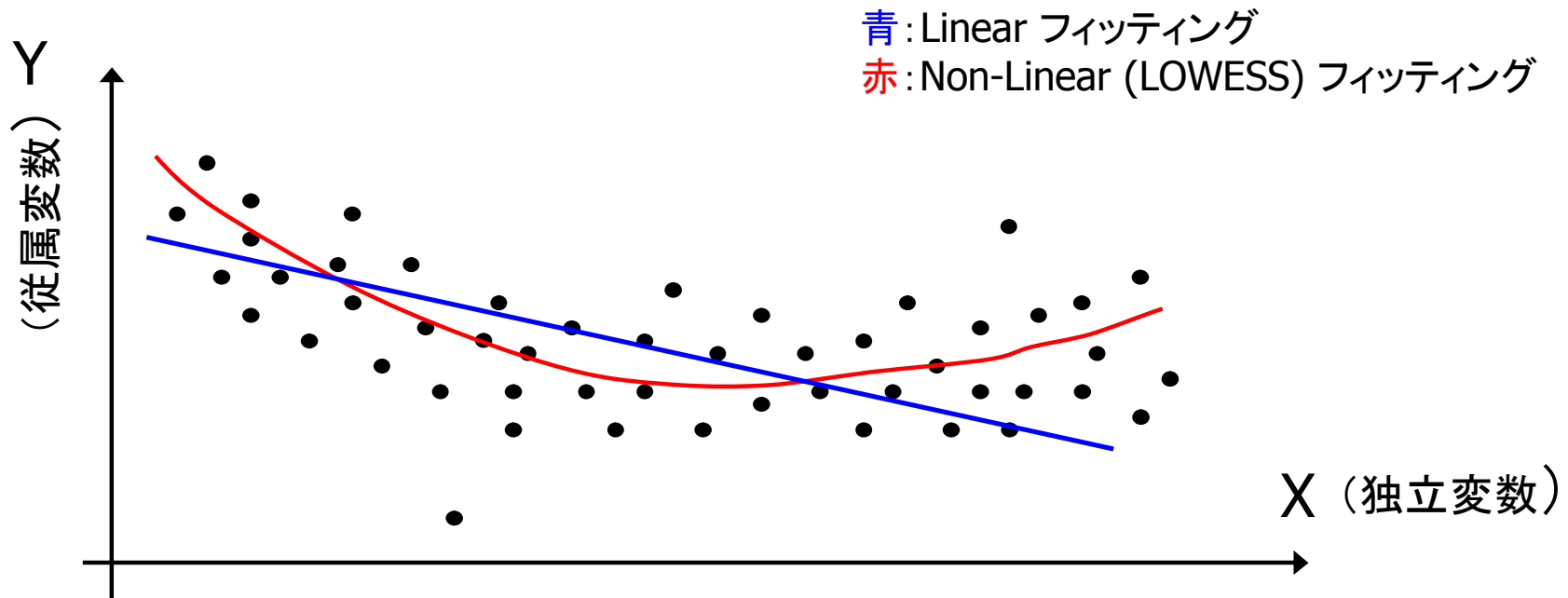
$$LOWESSDyeNormFactor = \frac{DyeNormSignal}{BGSubSignal}$$





Correct Dye Biases

Linear vs. Non-Linear フィットティング



Linear フィットティング: $y = \text{Slope (傾き)} * x + \text{Intercept (切片)} + \text{scatter}$

$$y = m * x + c$$

1次式での当てはめ
→ 線形近似 (i.e. Linear)

Linear フィットティングにおける仮定:

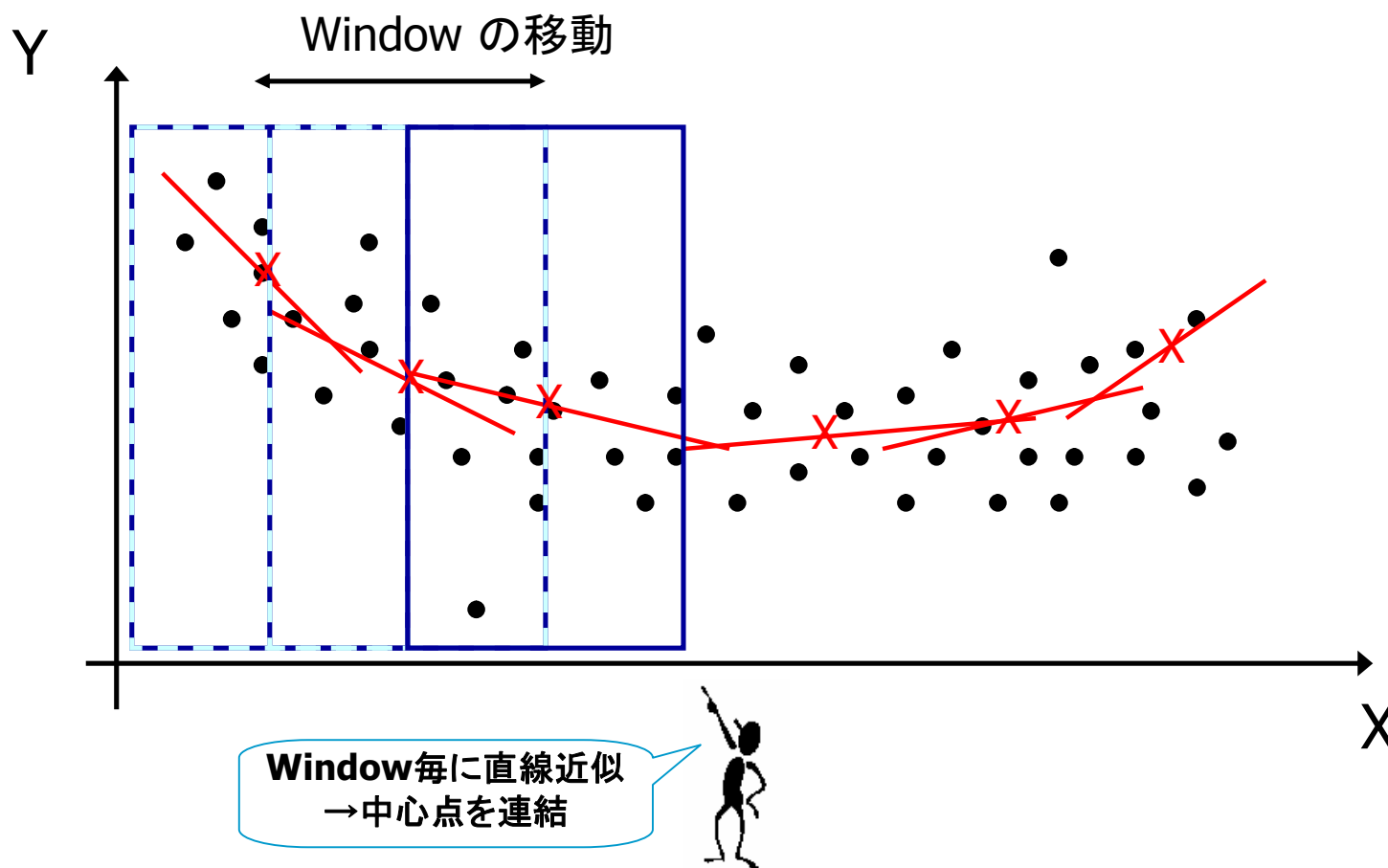
1. Scatter is Gaussian about a Mean = 0
2. Standard Deviation of scatter about a point on the curve is independent of the x-variable.





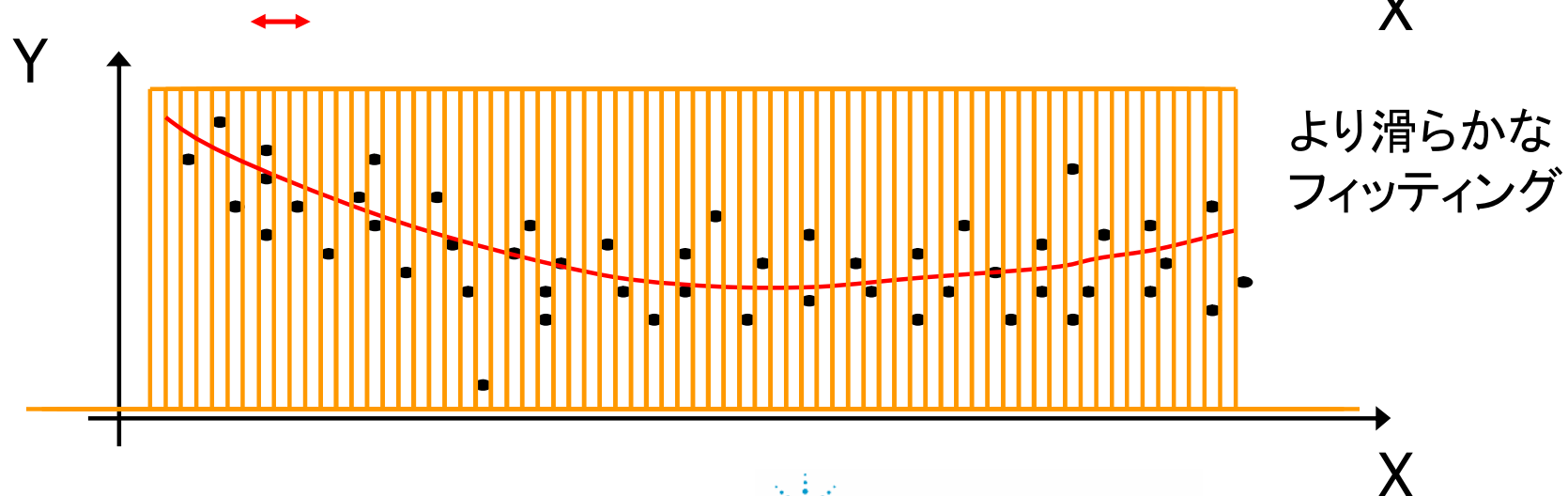
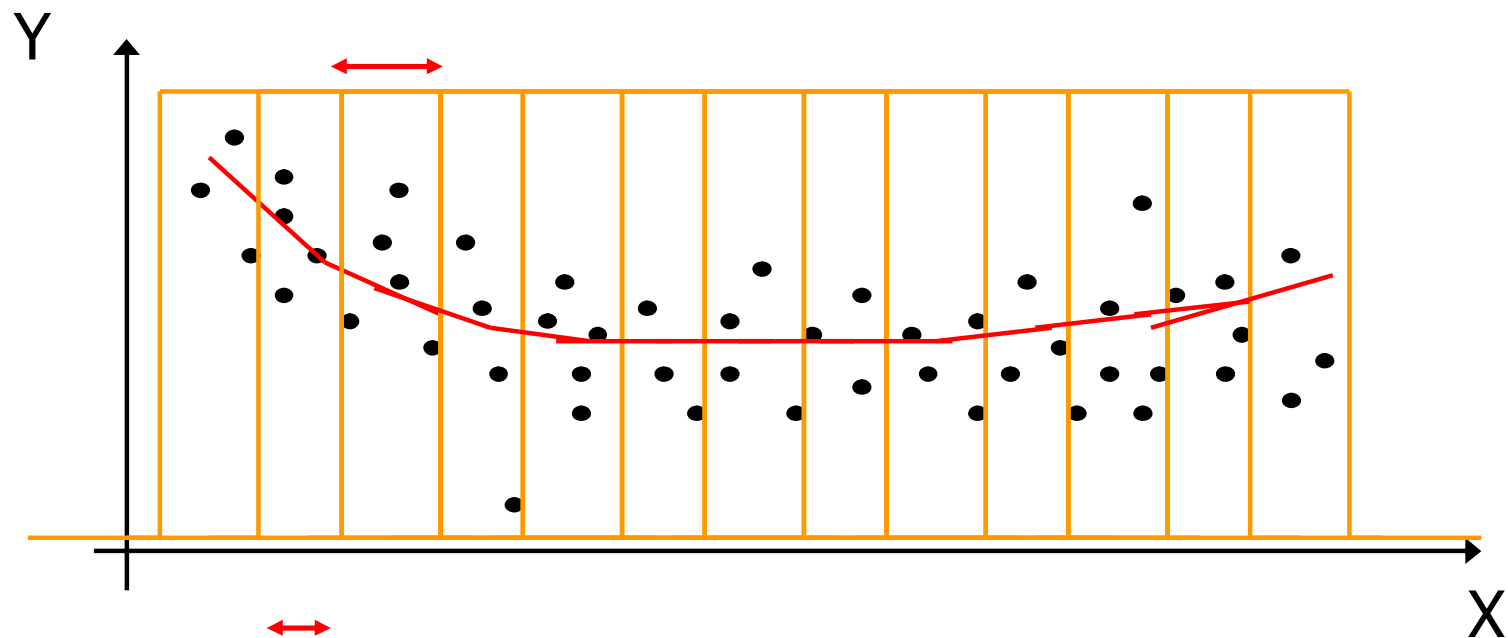
Correct Dye Biases LOWESS

Locally Weighted Linear Regression (局所重み付け直線回帰)





Correct Dye Biases LOWESSフィッティングの細分化





Correct Dye Biases **DyeNormSignal の計算**

赤・緑チャンネル、フィーチャ毎に色素補正後のシグナル強度を計算

$$DyeNormSignal = BGSubSignal \times DNF$$

この時

$$DNF = LinearDyeNormFactor \text{ (リニア法)}$$

$$DNF = Linear \ \& \ LOWESSDyeNormFactor \text{ (リニア \& LOWESS法)}$$

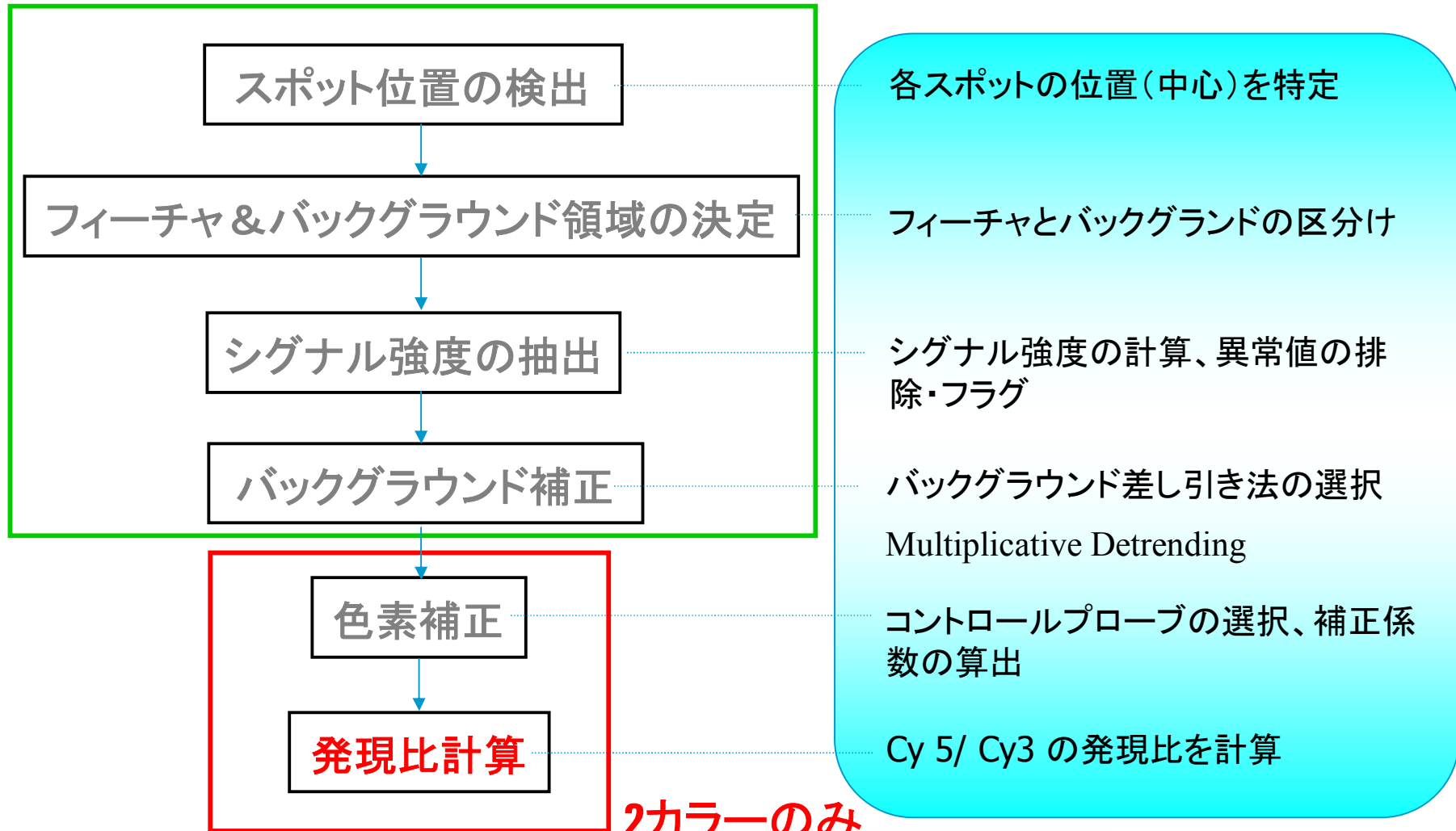
$$DNF = LOWESSDyeNormFactor \text{ (LOWESS法)}$$

いずれかの **DyeNormFactor (DNF)**を適用



数値化(スポット解析)の流れ

1、2カラー共通



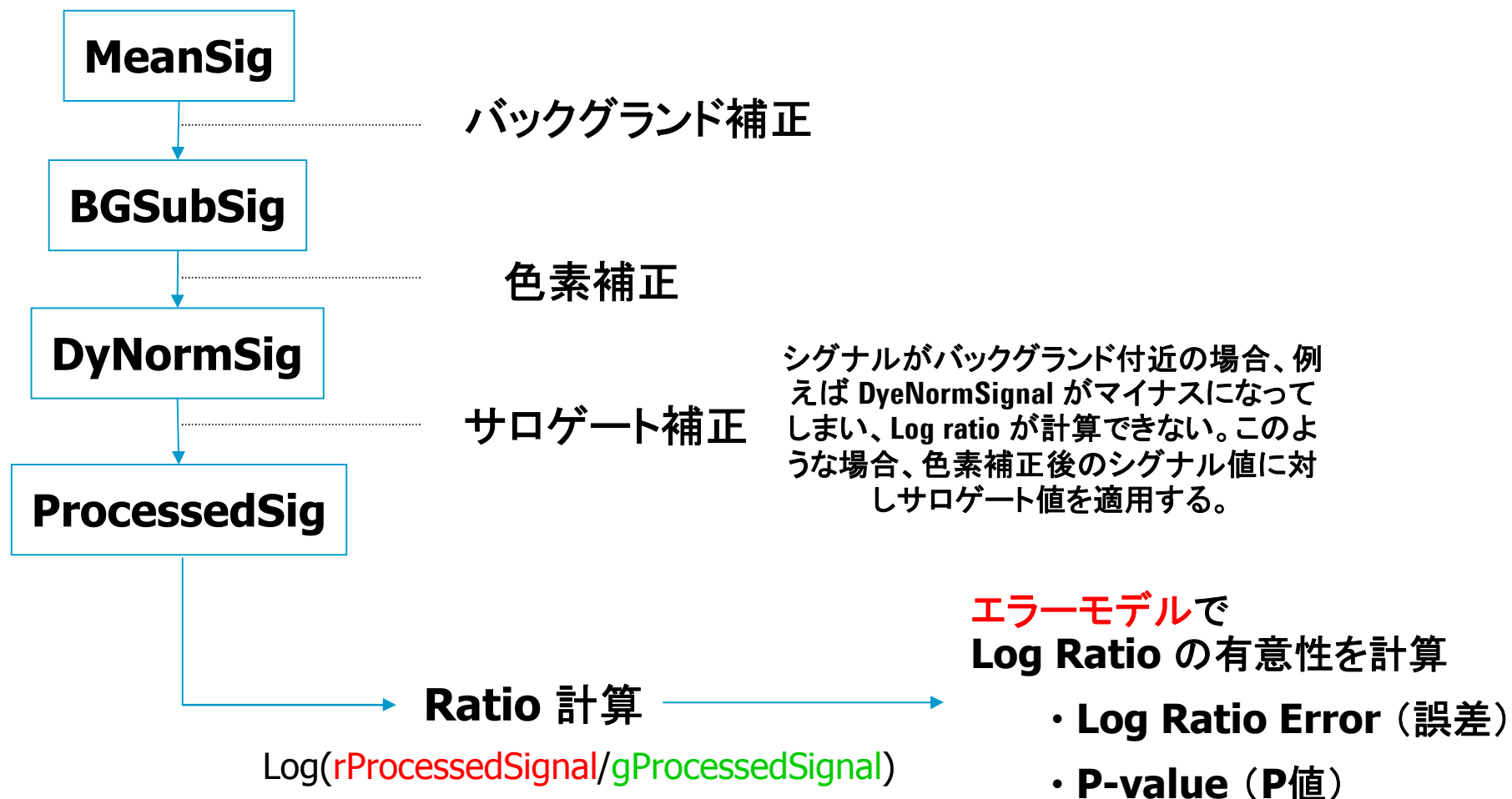
2カラーのみ





Compute Ratios

発現比の計算の流れ





Compute Ratios

P値および Log Ratio エラーの計算

式 1

$$pvalue = 1 - \text{Erf}\left(\frac{|xDev|}{\sqrt{2}}\right) = \text{Erfc}\left(\frac{|xDev|}{\sqrt{2}}\right)$$

Erf: 確率密度関数
正規分布 $N(0, 1/2)$ の2倍

P値は帰無仮説 (LogRatio=0)
が誤って棄却され得る統計的
可能性

式 2

$$xDev = \frac{\text{LogRatio}}{\text{LogRatioError}}$$

xDev: LogRatio=0からの偏差に相当。S/Nパラメータに類似。
FE output のFEATURES テーブルに記載。

ユニバーサルエラーモデルの場合
xDev から LogRatioError を計算

伝播エラーモデルの場合
LogRatioError から xDev を計算





Compute Ratios

エラーモデルのデフォルト設定変更

Protocol Steps	
Place Grid	
Find Spots	
Flag Outliers	
Compute Bkgd, Bias and Error	
Correct Dye Biases	
Compute Ratios	
Calculate Metrics	
Generate Results	
	Settings
	Background Subtraction Method
	No Background Subtraction
	☒ Significance (for IsPosAndSignif and IsWellAboveB
	Use Error Model for Significance
	2-sided t-test of feature vs. background maximum
	0.01000
	WellAboveMulti
	13.0
	☒ Signal Correction
	☒ Calculate Surface Fit (required for Spatial Detrend
	True
	Feature Set For Surface Fit
	FeaturesInNegativeControlRange
	Perform Filtering For Surface Fit
	True
	Perform Spatial Detrending
	True
	☒ Adjust Background Globally
	False
	☒ Perform Multiplicative Detrending
	True
	Detrend on Replicates Only
	True
	☒ Filter Low signal probes from Fit?
	True
	Neg. Ctrl. Threshold Mult. Detrend Factor
	5
	Perform Filtering for Fit
	Use Window Average
	Robust Neg Ctrl Stats?
	False
	☒ Choose universal error, or the most conservative (o
	Most Conservative
	MultErrorGreen
	0.1000
	MultErrorRed
	0.1000
	Auto Estimate Add Error Red
	True
	Auto Estimate Add Error Green
	True
	Use Surrogates
	True

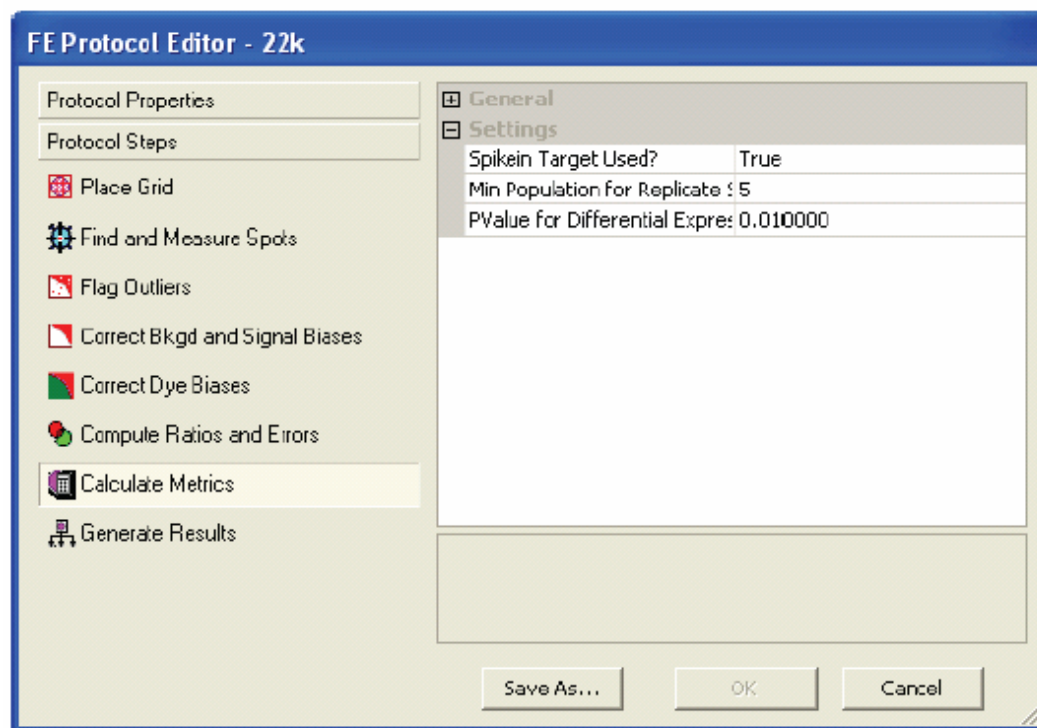
Additive エラーの自動見積もり (Agilentアレイのデフォルト) :

- ユニバーサルエラーにおける Additive エラー項を、Spatial Detrend Surface (遺伝子発現) または Negative Control (CGH, CoC) の変動に基づいて動的に算出
- Additive エラーの自動見積もりを行う際は、バックグラウンド補正の中で Spatial Detrend を実施する事を推奨



Calculate Metrics

QCレポートパラメータの設定



3つのパラメータを変更可能：

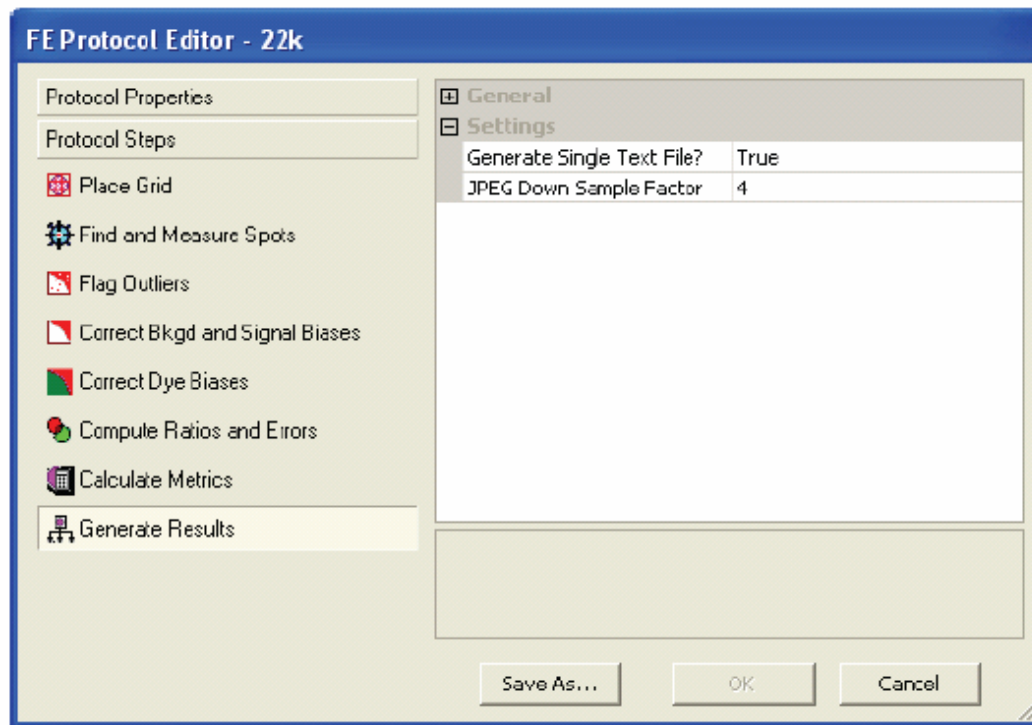
- SpikeIn Target Used
- Min Population for Replicates
- PValue for Differential Expression

Figure 50 FE Protocol Editor with Calculate Metrics selected



Generate Results

出力ファイル形式の設定



2つのパラメータを変更可能：

- Generate Single Text File
- JPEG Down Sampling Factor

Figure 51 FE Protocol Editor with Generate Results selected



QCLレポート

数値化ソフトウェア Feature Extraction 8.1 以降の機能 自動QCレポート出力

QCレポートの詳細は、

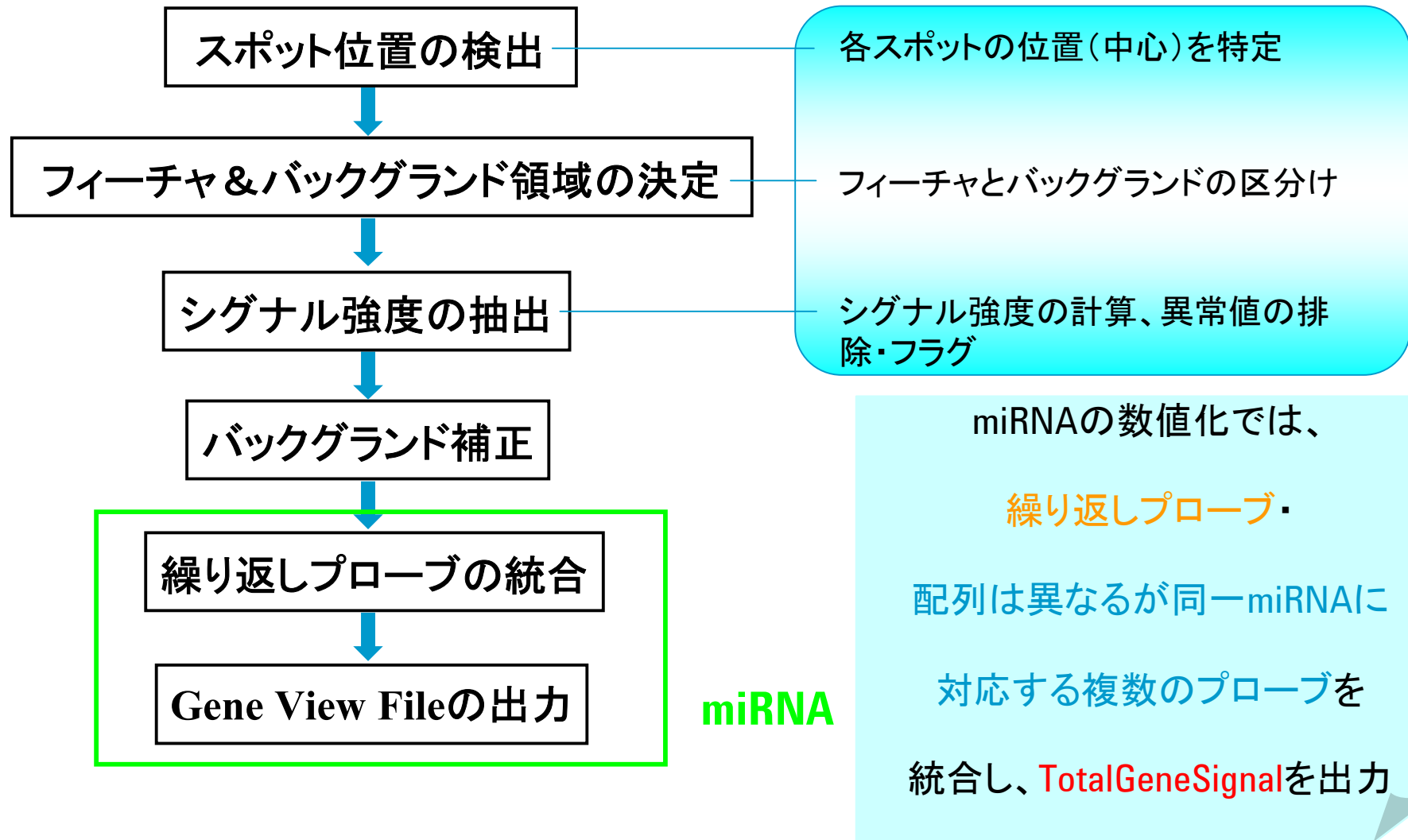
Help > Reference Guideを選択

Refernce Guide内のQC Report Resultsに記載



※日本語版ドキュメントも準備しております。

数値化(スポット解析)の流れ



GeneViewファイルの作成

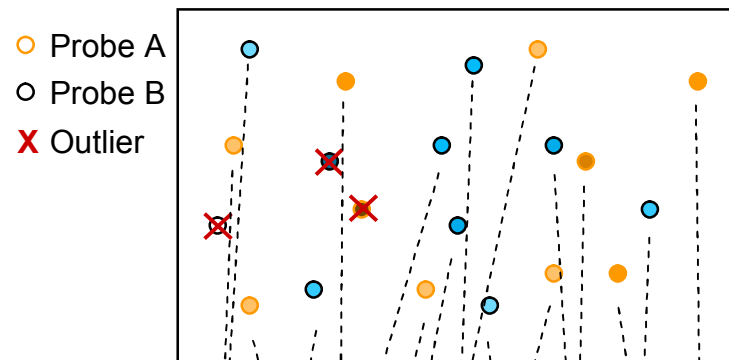
miRNA FE protocolでは、タブ区切りテキストファイル・MAGE-MLファイル・QCレポートなどに加え、**GeneView**という名前のタブ区切りテキストファイルを出力します。

SystematicName	プローブがハイブリするよう設計されたターゲット配列のID
ControlType	フィーチャーのコントロールタイプ 0 コントロール以外のプローブ 1 ポジティブコントロール -1 ネガティブコントロール
gTotalGeneSignal	各miRNAについて計算されたシグナル値
gTotalGeneError	シグナル強度のエラー値
gIsGeneDetected	シグナル値がエラー値よりも十分高く検出されているか 1 シグナル値がエラー値の3倍以上である 0 シグナル値がエラー値の3倍より低い

text	integer	float	float	boolean
SystematicName	ControlType	gTotalGeneSignal	gTotalGeneError	gIsGeneDetected
DarkCorner	1	4.89901	3.4459	0
NC1_00000197	-1	7.29536	8.43373	0
NC1_00000215	-1	1.18085	8.3791	0
NC2_00079215	-1	0.121326	8.38398	0
NC2_00092197	-1	-0.309608	8.37845	0

TotalGeneSignalの計算法

- 複数のProbe配列/miRNA
- 複数の繰り返しFeature/Probe配列



Step 1 –Population Outlierを除外

Step 2 –ProbeごとのNon-Outlier Featureのシグナル強度を平均する

Step 3 –上の平均値に、このProbeに対応するPixelの数と補正係数をかける
= **TotalProbeSignal**

Step 4 –それぞれのmiRNAに対応する全てのProbeのTotalProbeSignalを合計する= **TotalGeneSignal**

$$[(\text{orange circle} + \text{orange circle} + \text{orange circle} + \text{orange circle} + \text{orange circle} + \text{orange circle} + \text{orange circle} + \text{orange circle} + \text{orange circle} + \text{orange circle}) / 9] * 10 * (\# \text{pixels/feature}) = \text{TotalProbeSignal (Probe A)}$$

$$[(\text{blue circle} + \text{blue circle} + \text{blue circle} + \text{blue circle} + \text{blue circle} + \text{blue circle} + \text{blue circle} + \text{blue circle} + \text{blue circle} + \text{blue circle}) / 8] * 10 * (\# \text{pixels/feature}) = \text{TotalProbeSignal (Probe B)}$$

$$\text{TotalProbeSignal (Probe A)} + \text{TotalProbeSignal (Probe B)} = \text{TotalGeneSignal}$$

Feature ExtractionでのTotalGeneSignalの計算法例

例) 1 miRNAに対して2種類の配列 →それぞれの配列が10回繰り返し

Probeの種類	miRNAの種類	各Probeの Signal強度	Pop Outlier フラグ	Total Probe Signal	Total Gene Signal
ProbeName	SystematicName	gProcessedSignal	gIsFeatPopnOL	gTotalProbeSignal	gTotalGeneSignal
A_25_P00010486	hsa-miR-135b	53.76	0	126.73	126.73
A_25_P00011017	hsa-miR-136	346.94	0	230.453	680.864
A_25_P00011017	hsa-miR-136	330.65	0	230.453	680.864
A_25_P00011017	hsa-miR-136	330.81	0	230.453	680.864
A_25_P00011017	hsa-miR-136	335.42	0	230.453	680.864
A_25_P00011017	hsa-miR-136	342.90	0	230.453	680.864
A_25_P00011017	hsa-miR-136	349.93	0	230.453	680.864
A_25_P00011017	hsa-miR-136	338.75	0	230.453	680.864
A_25_P00011017	hsa-miR-136	345.95	0	230.453	680.864
A_25_P00011017	hsa-miR-136	345.28	0	230.453	680.864
A_25_P00011017	hsa-miR-136	369.77	1	230.453	680.864
A_25_P00011018	hsa-miR-136	668.70	0	450.411	680.864
A_25_P00011018	hsa-miR-136	681.78	0	450.411	680.864
A_25_P00011018	hsa-miR-136	643.20	0	450.411	680.864
A_25_P00011018	hsa-miR-136	690.24	0	450.411	680.864
A_25_P00011018	hsa-miR-136	686.91	0	450.411	680.864
A_25_P00011018	hsa-miR-136	659.15	0	450.411	680.864
A_25_P00011018	hsa-miR-136	663.02	0	450.411	680.864
A_25_P00011018	hsa-miR-136	649.57	0	450.411	680.864
A_25_P00011018	hsa-miR-136	657.28	0	450.411	680.864
A_25_P00011018	hsa-miR-136	659.71	0	450.411	680.864
A_25_P00010616	hsa-miR-137	333.21	0	226.551	443.472
A_25_P00010616	hsa-miR-137	326.11	0	226.551	443.472



Feature ExtractionでのTotalGeneSignalの計算法例

ProbeName	SystematicName	gProcessedSignal	glsFeatPopnOL
A_25_P00010486	hsa-miR-135b	53.76	0
A_25_P00011017	hsa-miR-136	346.94	0
A_25_P00011017	hsa-miR-136	330.65	0
A_25_P00011017	hsa-miR-136	330.81	0
A_25_P00011017	hsa-miR-136	335.42	0
A_25_P00011017	hsa-miR-136	342.90	0
A_25_P00011017	hsa-miR-136	349.93	0
A_25_P00011017	hsa-miR-136	338.75	0
A_25_P00011017	hsa-miR-136	345.95	0
A_25_P00011017	hsa-miR-136	345.28	0
A_25_P00011017	hsa-miR-136	369.77	1
A_25_P00011018	hsa-miR-136	668.70	0
A_25_P00011018	hsa-miR-136	681.78	0
A_25_P00011018	hsa-miR-136	643.20	0
A_25_P00011018	hsa-miR-136	690.24	0
A_25_P00011018	hsa-miR-136	686.91	0
A_25_P00011018	hsa-miR-136	659.15	0
A_25_P00011018	hsa-miR-136	663.02	0
A_25_P00011018	hsa-miR-136	649.57	0
A_25_P00011018	hsa-miR-136	657.28	0
A_25_P00011018	hsa-miR-136	659.71	0
A_25_P00010616	hsa-miR-137	333.21	0
A_25_P00010616	hsa-miR-137	326.11	0

①PopOL
は除外

②残り9プ
ローブの平均
値を算出

1. 繰り返しプローブのうち、
Population Outlierフラグの
立っているものを除外する
2. 残りの繰り返しプローブの
Processed Signalの平均値
を出す

Feature ExtractionでのTotalGeneSignalの計算法例

Font		Alignment	
f _x		Merge & Center	
	BT		
	float		integer
tage	SpotAnalysis_ExclusionZonePercentage		SpotAnalysis_Esti
	0.7		1.2
	float		float
fference	EffectiveFeatureSizeFraction		FeatureUniformit
	0		0.611459

3. EffectiveFeatureSizeFraction

(スポットの中の、シグナル抽出計算

に用いられたエリアの割合)を算出

※値はTextOutputFileのStatsに報告されます(左図参照)

4. スポットの面積値を算出

※8x15kの場合は

面積値 = (スポット半径 = 32.5) の2乗 × 3.14

5. 以下の式で、TotalProbeSignalを算出する

TotalProbeSignal=

PopOLフラグの立っていないプローブの平均シグナル強度

× 繰り返しプローブ数

× Effective Feature Fraction

× スポット面積値

× 1/30000

Feature ExtractionでのTotalGeneSignalの計算法例

6.それぞれのProbeについてTotalProbeSignalを算出し、同じ

miRNAに対応するTotalProbeSignalを合計する=TotalGeneSignal

ProbeName	SystematicName	gProcessedSignal	gIsFeatPopnOL	gTotalProbeSignal	gTotalGeneSignal
A_25_P00010486	hsa-miR-135b	53.76	0	126.73	126.73
A_25_P00011017	hsa-miR-136	346.94	0	230.453	680.864
A_25_P00011017	hsa-miR-136	330.65	0	230.453	680.864
A_25_P00011017	hsa-miR-136	330.81	0	230.453	680.864
A_25_P00011017	hsa-miR-136	335.42	0	230.453	680.864
A_25_P00011017	hsa-miR-136	342.90	0	230.453	680.864
A_25_P00011017	hsa-miR-136	349.93	0	230.453	680.864
A_25_P00011017	hsa-miR-136	338.75	0	230.453	680.864
A_25_P00011017	hsa-miR-136	345.95	0	230.453	680.864
A_25_P00011017	hsa-miR-136	345.28	0	230.453	680.864
A_25_P00011017	hsa-miR-136	369.77	1	230.453	680.864
A_25_P00011018	hsa-miR-136	668.70	0	450.411	680.864
A_25_P00011018	hsa-miR-136	681.78	0	450.411	680.864
A_25_P00011018	hsa-miR-136	643.20	0	450.411	680.864
A_25_P00011018	hsa-miR-136	690.24	0	450.411	680.864
A_25_P00011018	hsa-miR-136	686.91	0	450.411	680.864
A_25_P00011018	hsa-miR-136	659.15	0	450.411	680.864
A_25_P00011018	hsa-miR-136	663.02	0	450.411	680.864
A_25_P00011018	hsa-miR-136	649.57	0	450.411	680.864
A_25_P00011018	hsa-miR-136	657.28	0	450.411	680.864
A_25_P00011018	hsa-miR-136	659.71	0	450.411	680.864
A_25_P00010616	hsa-miR-137	333.21	0	226.551	443.472
A_25_P00010616	hsa-miR-137	326.11	0	226.551	443.472

Protocols for Feature Extraction 9.5

アプリケーション	FE Protocol	Agilent アレイ フォーマット	スキャナ	実験プロトコル
CGH	CGH-4_95_Feb07	2x11K, 22K, 44K, 8x15K, 4x44K, 2x105K, 244K	Agilent, GenePix	アレイCGH v4.0 (Publication Number G4140-90010)
	CGH-4_95_Tecan_Feb07	2x105K, 244K	Agilent, GenePix	アレイCGH Tecan HS Pro Hybridizationを使う場合 (P/N G4410-90011)
Chip-on-Chip	ChIP-v1_95_May07	4x44K, 2x105K, 244K	Agilent, GenePix	Mammalian Chip-on-Chip (P/N G4481-90010)
遺伝子発現 (1色法)	GE1-v5_95_Feb07	8x15K, 4x44K, 2x105K, 244K	Agilent, GenePix	1色法遺伝子発現 v5.0.xあるいはv5.5.x (高密度フォーマット対応) (P/N G4140-90040)
	GE1-v1_95_Feb07	2x11K, 22K, 44K	Agilent, GenePix	1色法遺伝子発現 v1.0.x (P/N G4140-90040)
遺伝子発現 (2色法)	GE2-v5_95_Feb07	8x15K, 4x44K, 2x105K, 244K	Agilent, GenePix	2色法遺伝子発現 v5.0.xあるいはv5.5.x (高密度フォーマット対応) (P/N G4140-90050)
	GE2-v4_95_Feb07	2x11K, 22K, 44K	Agilent, GenePix	2色法遺伝子発現 v4.0.x (P/N G140-90050)
	GE2-SSPE_95_Feb07	2x11K, 22K, 44K	Agilent, GenePix	2色法遺伝子発現 v3.0.x (SSPEまたはSSC洗浄) (P/N G4140-90050または G4140-90030)
	GE2-NonAT_95_Feb07	Non-Agilent	Agilent	-
microRNA	miRNA-v1_95_May07	8x15K	Agilent	microRNA v1.0 (P/N G4170-90010)



Protocols for Feature Extraction 9.1

アプリケーション	FE Protocol	Agilent アレイ フォーマット	スキャナ	実験プロトコル
CGH	CGH-v4_91	2x11K, 22K, 44K, 2x105K, 244K	Agilent, GenePix	アレイCGH v4.0 (P/N G4140-90010)
遺伝子発現(1色法)	GE1-v5_91_0806	4x44K, 2x105K, 244K	Agilent, GenePix	1色法遺伝子発現 v5.0.x (高密度フォーマット対応) (P/N G4140-90040)
	GE1-v1_91	2x11K, 22K, 44K	Agilent, GenePix	1色法遺伝子発現 v1.0.x (P/N G4140-90040)
遺伝子発現(2色法)	GE2-v5_91_0806	4x44K, 2x105K, 244K	Agilent, GenePix	2色法遺伝子発現 v5.0.x (高密度フォーマット対応) (P/N G4140-90050)
	GE2-v4_91	2x11K, 22K, 44K	Agilent, GenePix	2色法遺伝子発現 v4.0.x (P/N G140-90050)
	GE2-SSPE_91	2x11K, 22K, 44K	Agilent, GenePix	2色法遺伝子発現 v3.0.x (SSPEまたはSSC洗浄) (P/N G4140-90050またはG4140-90030)
	GE2- NonAT_91_0806	Non-Agilent	Agilent	-