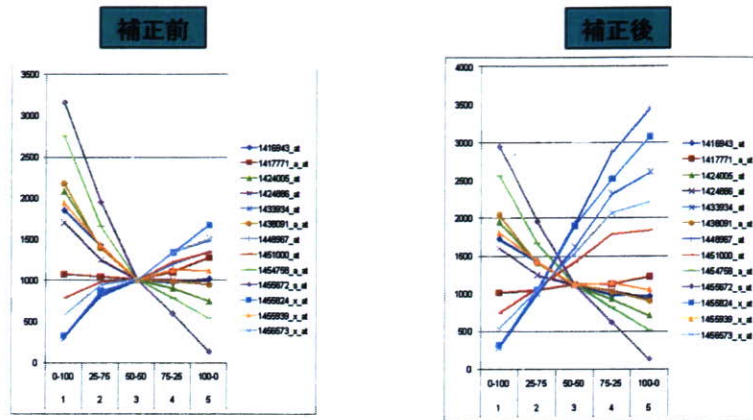


1.8.5. 全プローブ計算④検証① 高値補正付き



- 対象プローブセット: 50%:50%において、1000程度の値を持つ



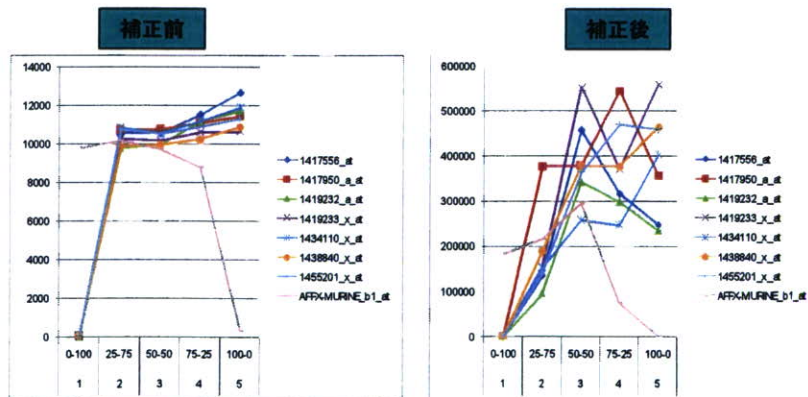
高値を補正する式にすると若干安定する

Copyright © NTT COMWARE 2008

1.8.5. 全プローブ計算④検証①



- 対象プローブセット: 50%:50%において、10000以上の値を持つ



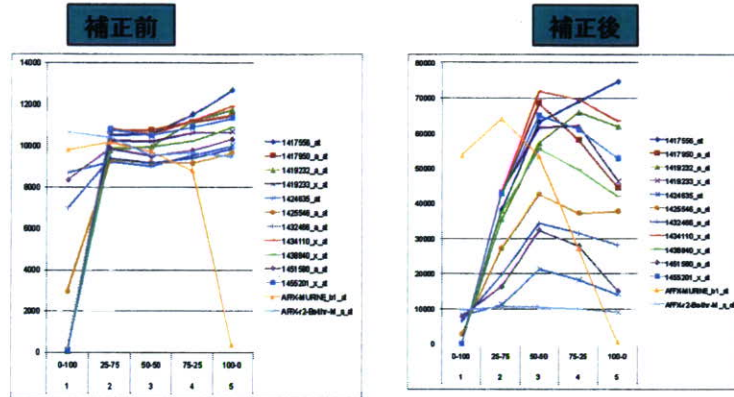
0%から50%まではすこし上向きだが直線に見えなくもない。しかし、50%を超えると、崩れている。Lngmuirの最大値推定に問題がありそうである

Copyright © NTT COMWARE 2008

1.8.5. 全プローブ計算④検証① 高値補正付き



- 対象プローブセット: 50%:50%において、10000以上の値を持つ



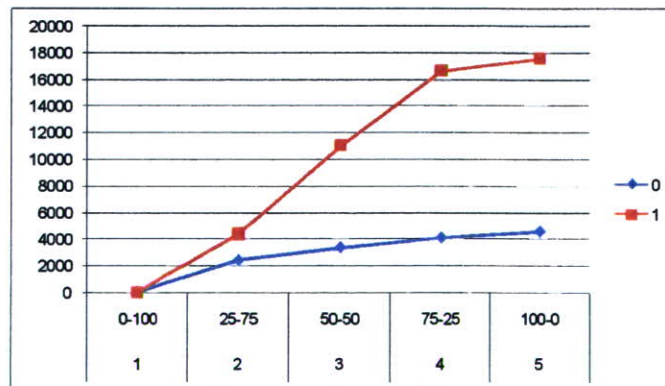
高値に対する補正を行う変形をすると、若干安定する

Copyright © NTT COMWARE 2008

1.8.5. 全プローブ計算④検証① 高値補正付



- Cypla2の結果



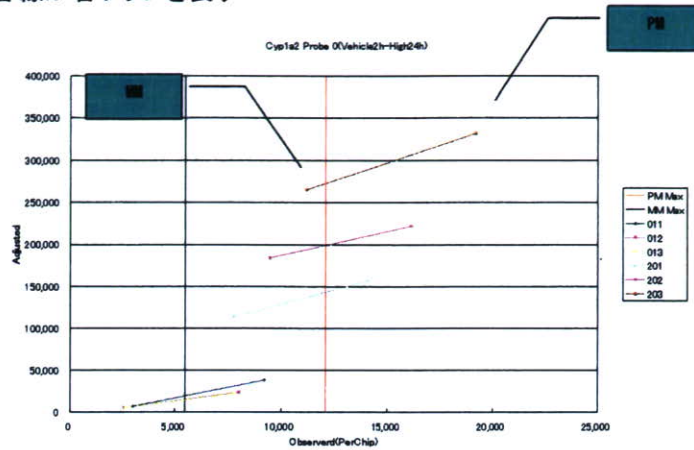
高値に対する補正を行う変形をすると、若干安定する

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



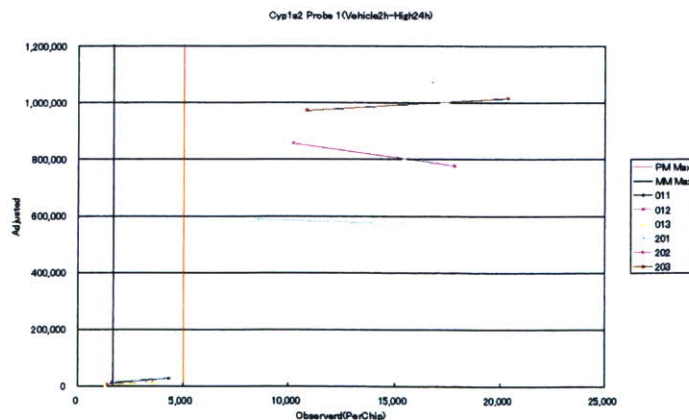
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



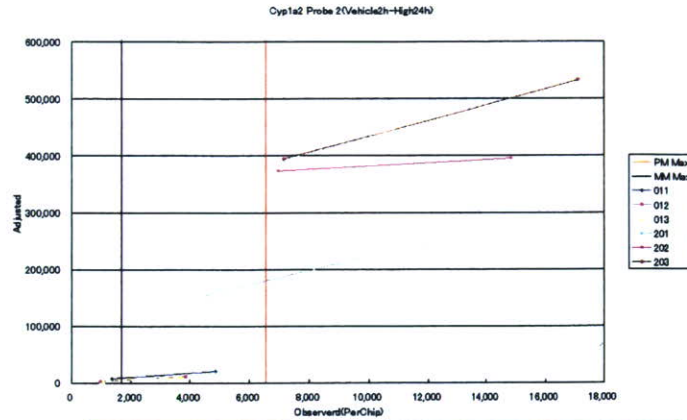
計測値はLangmuir曲線の最大値を大きく上回っている。
補正後にPM/MMの大小関係が逆転しているペアが存在する

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



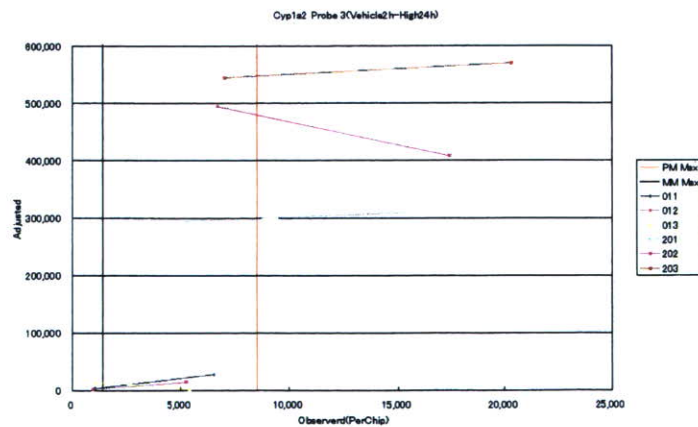
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



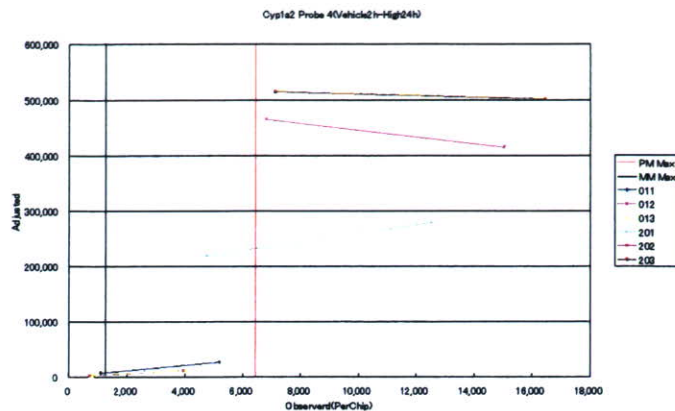
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



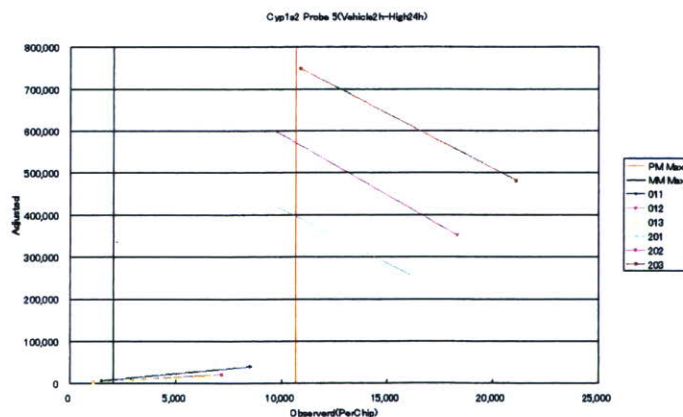
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



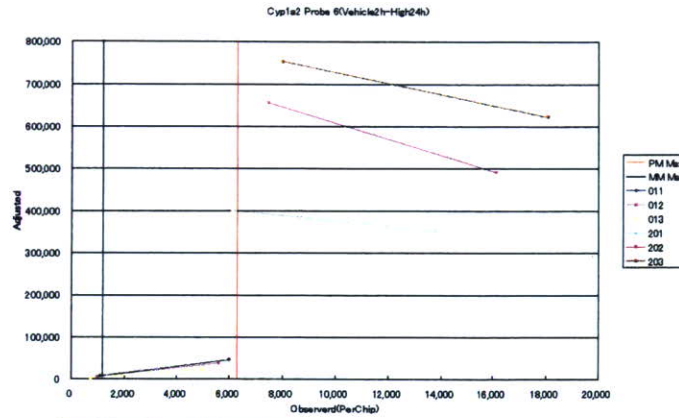
PM-MMで値が逆転しており、結果を不安定にしている

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



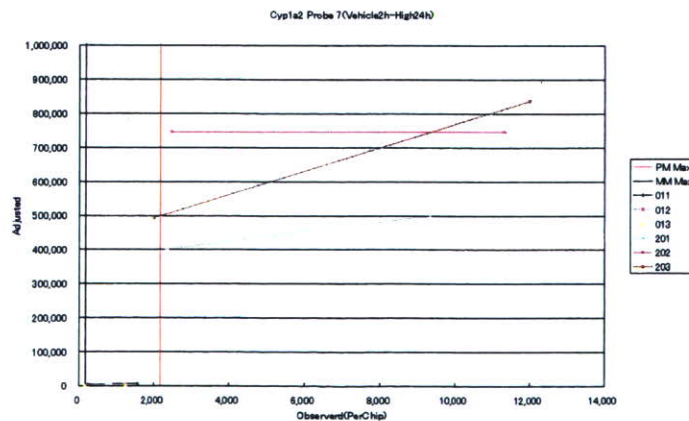
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



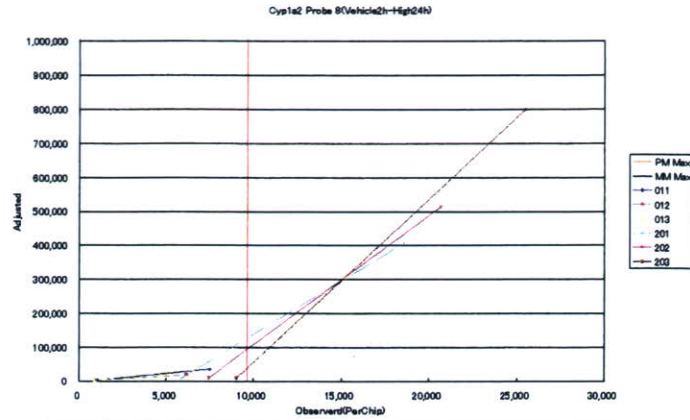
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



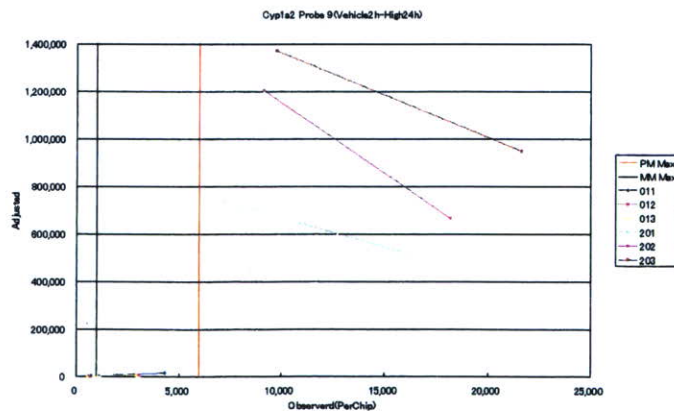
MMはLangmuirが選択されていない

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



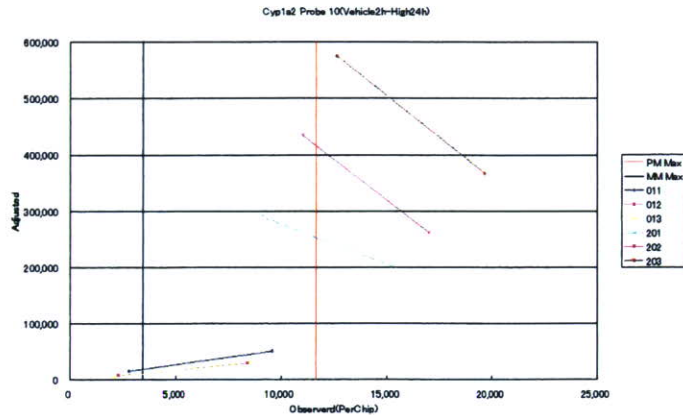
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.6. 全プローブ計算④検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.8.7. 全プローブ計算④課題: 変換範囲



- Cyp1a2において、変換結果として良好な形状を示さなかった。
- 考察

Liverサンプルであり、TCDDの影響により通常状態より大きな値を示している



LBMIにおけるRNAの濃度範囲を逸脱している



本来のLangmuir方程式より低い値で飽和しているような値が推定されている



誤差を対数領域で計算する(幾何平均)と
線形領域で計算する(算術平均)よりも小さな値を示す

Copyright © NTT COMWARE 2008

1.9. 全プローブ計算⑤



- 濃度平衡係数がPM/MMIにおいて同じであると仮定する場合と仮定しない場合で違いがあるのか？
- アイデア
 - PM/MMIにおいて対象とするRNAの配列は同じものである。同じRNAを対象としているので、同一チップでは、同一の濃度であるとみなすことが可能なはずである。

Copyright © NTT COMWARE 2008

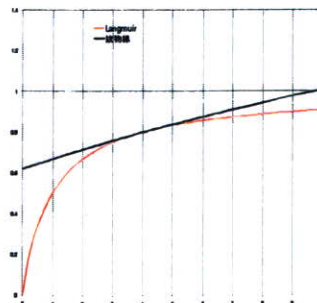
1.9.1. 全プローブ計算⑤数値計算上の技巧



- 数値計算を行う上で、Langmuirの方程式は上に有界な数式である。次の2点で数値計算上問題がある。
 - 濃度に対する係数に鈍感な関数で、完全飽和に近い状態で適切な収束が得られない
 - 飽和していない場合には、AICにより、排除される
 - 今回は逆関数にするので、逆関数の際に右に有界で定義域が制限される。
- 高濃度において、平方根にフィットするものとみなす。
 - この領域に入った場合には詳細な検討が必要である。

$$f(x) = \begin{cases} \frac{x}{x+1} & \text{for } x \leq 4 \\ \frac{8}{\sqrt{125}\sqrt{x+6}} & \text{for } x \geq 4 \end{cases}$$

$$I_{\#} = \begin{cases} I_p \frac{k_p c_j}{k_p c_j + 1} + b g_p & \text{for } k_p c_j \leq 4 \\ I_p \sqrt{\frac{8}{125}} \sqrt{k_p c_j + 6} + b g_p & \text{for } k_p c_j \geq 4 \end{cases}$$



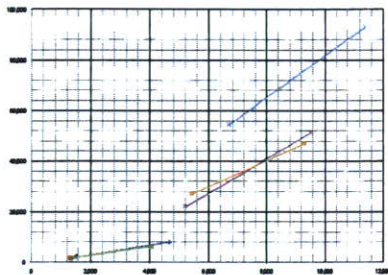
Copyright © NTT COMWARE 2008

1.9.2. 全プローブ計算⑤結果

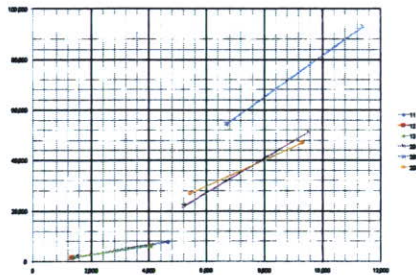


濃度平衡係数がPM/MMにおいて同じであると仮定する場合と仮定しない場合(Cyp1a2 Probe①)

PM/MM濃度連動モデル



PM/MM濃度独立モデル



係数は若干異なるが実験結果に大きな影響はない

Copyright © NTT COMWARE 2008

1.10. 全プローブ計算⑥



- 最大値を固定する方法
 - 最大値をLBMだけではなく、別の実験から求めて使用する
- アイデア
 - 最大値は、次の欠点が考えられる
 - 観測誤差の影響を受けやすい。
 - 似た別の配列が存在する場合に、Cross-Hybridizationにより、Langmuir方程式の仮定を崩す状況が考えられる
 - 現在まで非常に多くの実験を行ってきたこの中で最大値を飽和値とみなす。
 - 今回は、TTG20の最大値を飽和値の95%とみなす
 - 将来的には、全チップの上位10%の+5σを最大値とみなすようなことを考える

Copyright © NTT COMWARE 2008

1. 10. 1. 全プローブ計算⑥ Teradataによる線形領域モデル計算を用いた絞り込み

- Teradataを用いた線形計算を用いて、対象プローブの絞り込みを実施
- 次の条件を用いた
 - 上に凸(線形領域2次モデルの2次係数が負)
 - 線形領域線形モデルの傾きが平均値の5%以上
 - 線形領域線形モデルの傾きが標準偏差の5%以上

終了状態	Brain側飽和	Liver側飽和
勾配ベクトル判定収束	51,278	51,239
移動量判定収束	48,288	18,678
局所的最小値	50,203	50,139
総計	149,769	120,056

Copyright © NTT COMWARE 2008

1. 10. 2. 全プローブ計算⑥非線形最適化の課題&解決テクニック

- いくつかのプローブにおいては、適切な解に収束しなかった
 - 繰り返しの範囲量を超えた。または収束回数を超えた。
 - 非線形計算のパラメータを変更してやり直すようにプログラムに組み込み、ほとんどのプローブで収束するようにした
 - 繰り返しの最中に無限の数値が現れ、バッチモードのRが中断した
 - メッセージ「nlm」により有限でない値が与えられました」
 - 手作業でパラメータ変更実施

補正方法	飽和側	モデル	ProbesetNo	Pm/Mm	Probe	対処方法
PM	Liver	Langmuir	10709	MM	6	I00=18.5, BP0=16
	Liver	Linear	43161	PM	3	BP0を半分に
	Brain	Linear	1680	PM	10	BP0を半分に
	Brain	Linear	35906	PM	5	BP0を半分に
PMMM	Liver	Langmuir	11913	pm	4	I00=92, k00=1.11, BP0=11.9
	Brain	Linear	15256	pm	4	BP0を半分に

今後も発生する可能性はあるが、既に出来上がったモデルのパラメータを与えることで、回避できる。

Copyright © NTT COMWARE 2008

1.10.3. 全プローブ計算⑥非線形最適化の結果



- Liver側飽和では多くのプローブでLangmuirモデルが採用された。
- Brain側飽和では、Langmuirモデルが採用されたプローブは限られている

終了状態	Brain側飽和	B側採用個数	Liver側飽和	L側採用個数
勾配ベクトル判定収束	51,278	3	51,239	6,383
移動量判定収束	48,288		18,678	746
局所的最小値	50,203	16	50,139	2,742
総計	149,769	19	120,056	9,871

Copyright © NTT COMWARE 2008

1.10.4. 全プローブ計算⑥ 結果Langmuir Model採用個数



- ProbeSet中で何個のProbeがLangmuirモデルを採用しているか
- Brain側飽和

合計 / Count(*)	列ラベル					
行ラベル	0	1	2	3	4	総計
0		3	3	2		8
1	2					2
5					1	1
総計	2	3	3	2	1	11

Brain側飽和では非常に少ない

Copyright © NTT COMWARE 2008

1. 10. 4. 全プローブ計算⑥ Langmuir Model採用個数

• Liver側飽和

合計 / Count(*) 行ラベル	列ラベル	0	1	2	3	4	5	6	7	8	9	10	11	総計	
0			1089	225	124	81	42	41	17	10	6	3		1838	
1		1077	196	62	45	32	20	34	34	21	6	6		1533	
2			39	19	41	27	25	17	19	23	12	10	7	1	240
3			11	13	15	9	8	15	17	15	16	21	7	10	157
4			2	3	2	2	14	7	9	17	19	22	13	6	116
5						1	5	4	8	7	16	13	15	13	82
6						3	2	1	3	6	7	10	9	19	60
7			1					2		6	12	13	18	17	69
8					1	1	1	4	1	1	10	12	13	44	
9							1	1		1	3	7	7	20	
10								1	1		2	10	18	32	
11											1	5	5	11	
総計		1130	1320	345	212	188	110	137	127	115	117	112	109	4002	

PM側のLangmuirモデル採用個数が多いProbeSetが補正の影響を受けるはずである

Copyright © NTT COMWARE 2008

1. 10. 5. 全プローブ計算⑥検証①

• LBMでの線形性チェック

- 50:50を1に正規化したグラフを作成し、線形性のチェック
- Langmuir変換を含むプローブセットを全プロットし、目視で線形性の確認する。値の低いもの以外で変な値を示すものがないことを確認する

Copyright © NTT COMWARE 2008

1. 10. 5. 全プローブ計算⑥検証①



- 対象プローブセット: 50%:50%において、1000程度の値を持つ

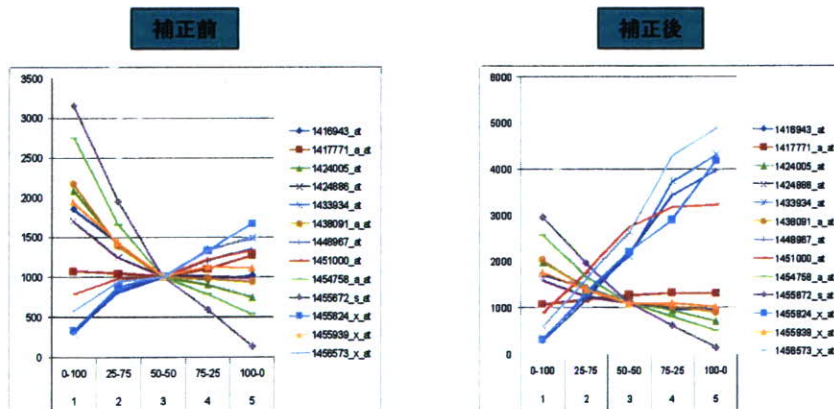
affy_id	Intensity
1424886_at	1008.53
1433934_at	1004.20
1424005_at	1002.17
1417771_a_at	1005.90
1451000_at	1000.37
1456573_x_at	1005.87
1455824_x_at	1006.07
1416943_at	1007.17
1454758_a_at	1007.60
1438091_a_at	1008.33
1448967_at	1003.97
1455939_x_at	1006.53
1455672_s_at	1004.00

Copyright © NTT COMWARE 2008

1. 10. 5. 全プローブ計算⑥検証①



- 対象プローブセット: 50%:50%において、1000程度の値を持つ



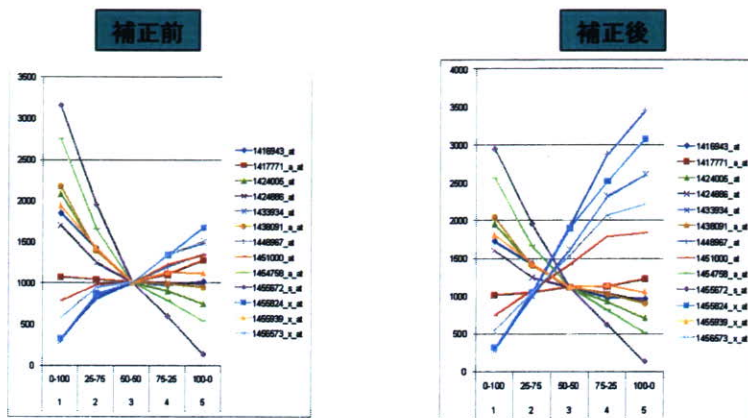
Liver 創発和のプローブは直線に引き延ばされ、高い値で、不安定になっている。

Copyright © NTT COMWARE 2008

1. 10. 5. 全プローブ計算⑥検証① 高値補正付き



- 対象プローブセット: 50%:50%において、1000程度の値を持つ



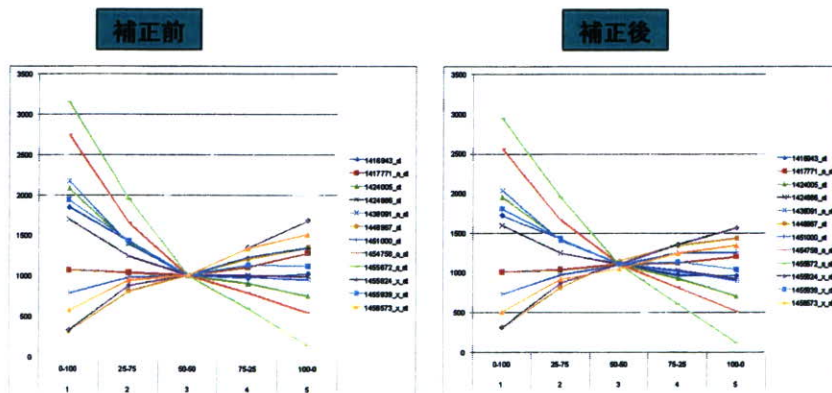
高値を補正する式にすると若干安定する

Copyright © NTT COMWARE 2008

1. 10. 5. 全プローブ計算⑥検証① 飽和値指定



- 対象プローブセット: 50%:50%において、1000程度の値を持つ



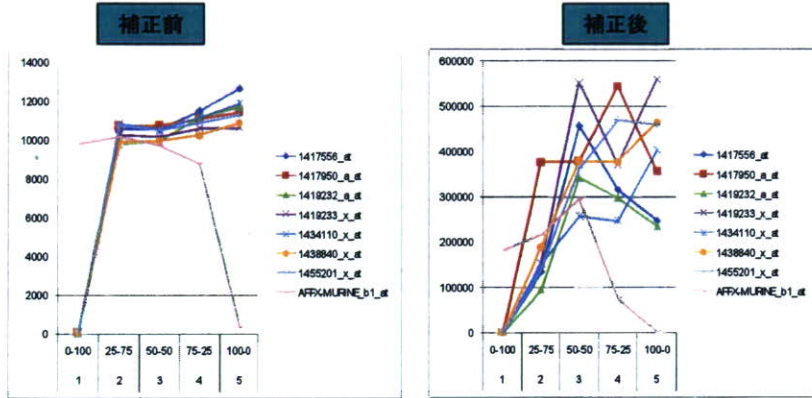
飽和値を指定すると安定する

Copyright © NTT COMWARE 2008

1.10.5. 全プローブ計算⑥検証①



- 対象プローブセット: 50%:50%において、10000以上の値を持つ



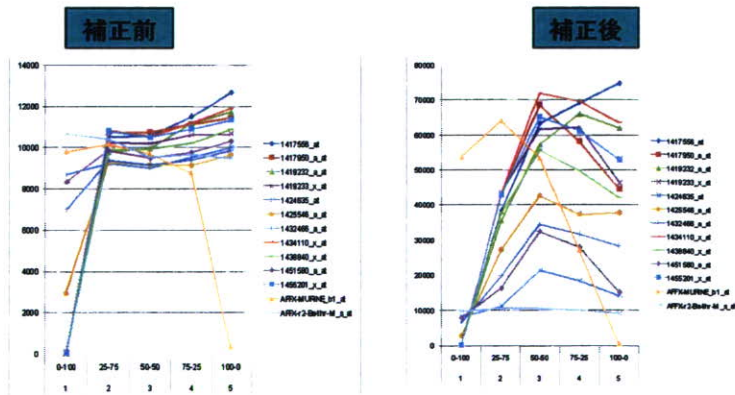
0%から50%まではすこし上向きだが直線に見えなくもない。しかし、50%を超えると、崩れている。Lnguirの最大値推定に問題がありそうである

Copyright © NTT COMWARE 2008

1.10.5. 全プローブ計算⑥検証① 高値補正付き



- 対象プローブセット: 50%:50%において、10000以上の値を持つ



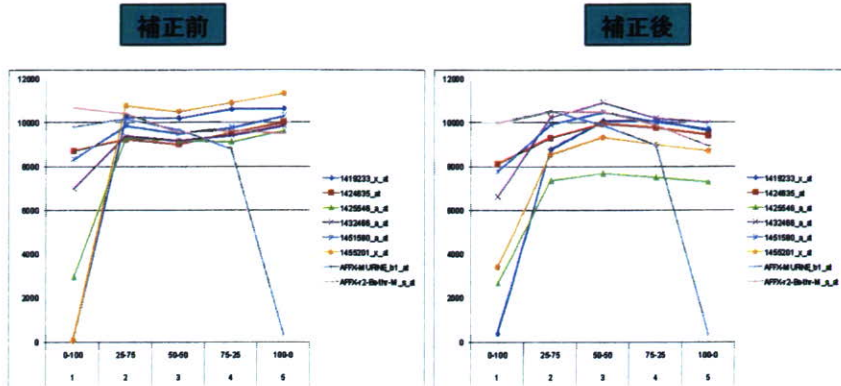
高値に対する補正を行う変形をすると、若干安定する

Copyright © NTT COMWARE 2008

1.10.5. 全プローブ計算⑥検証① 高値補正付き



- 対象プローブセット: 50%:50%において、10000以上の値を持つ



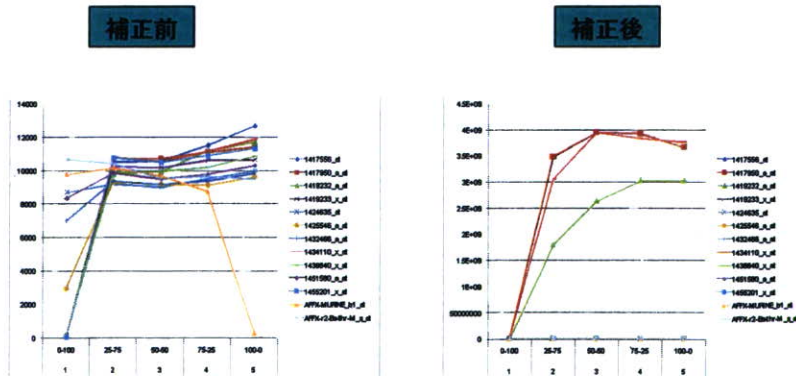
高値に対する補正を行う変形をすると、若干安定する

Copyright © NTT COMWARE 2008

1.10.5. 全プローブ計算⑥検証① 飽和値指定



- 対象プローブセット: 50%:50%において、10000以上の値を持つ



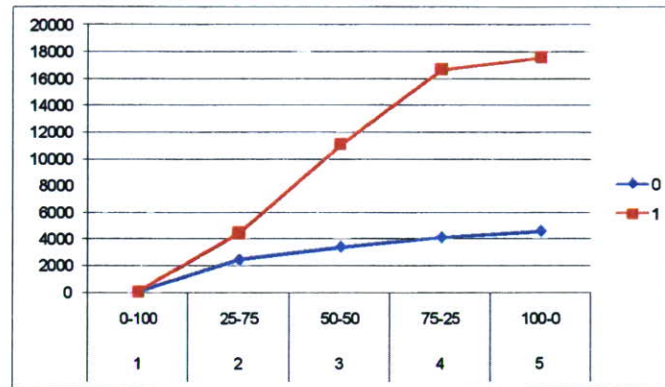
非線形最適化で失敗している。丹念に調べる必要がある

Copyright © NTT COMWARE 2008

1. 10. 5. 全プローブ計算⑥検証①



- Cyp1a2(1450715_at)



Copyright © NTT COMWARE 2008

1. 10. 6. 全プローブ計算⑥検証②



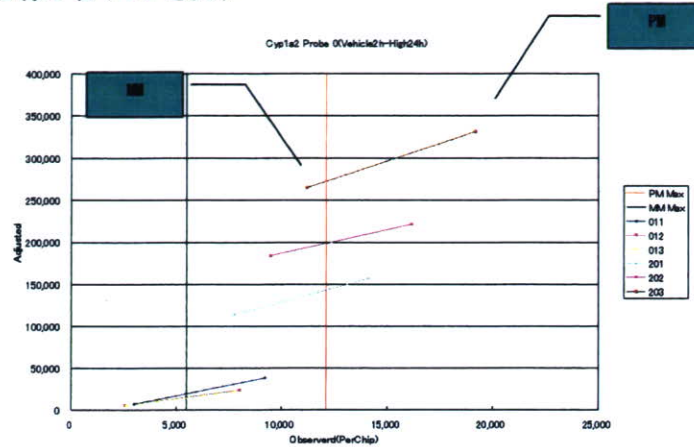
- 基本的確認
 - Percellome論文(BMC Genomics 2006, 7:64 doi:10.1186/1471-2164-7-64)において、QPCR計測を行っている遺伝子で飽和していると考えられる遺伝子が、本方法と、結果が合致しているかを確認する。
- 課題
 - QPCR対象のプローブセットがLangmuirモデルになっているか？
 - Cyp7a1は、Langmuirモデル対象となっている。
 - サーフェースまで描いた遺伝子はLangmuir対象は他になし。
 - 計測済みでは、他にあるか？
- 懸念事項
 - Percellome自身が飽和の影響を受けていて、高発現域でBiasを発生していないか？
- 対象データ
 - QPCRと同じ化合物を用いた実験のCELファイル
 - TTG20を使用

Copyright © NTT COMWARE 2008

1.10.6. 全プローブ計算⑥検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



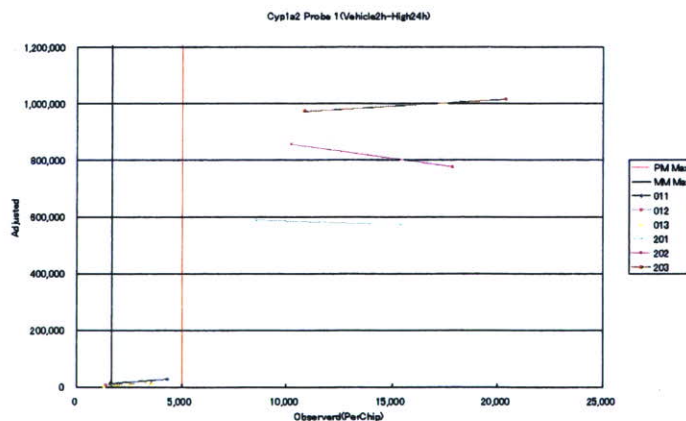
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.10.6. 全プローブ計算⑥検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



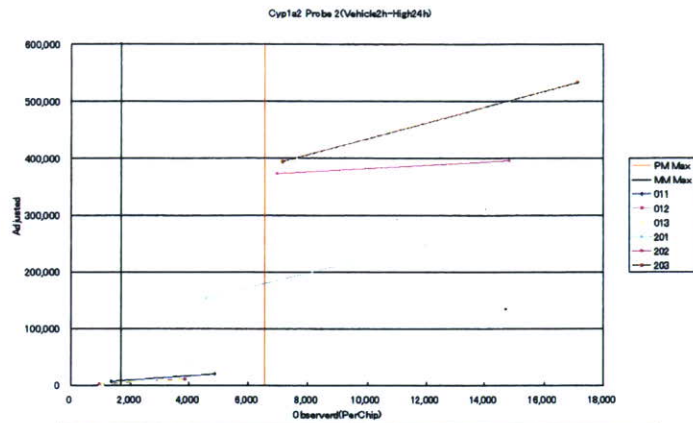
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.10.6. 全プローブ計算⑥検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



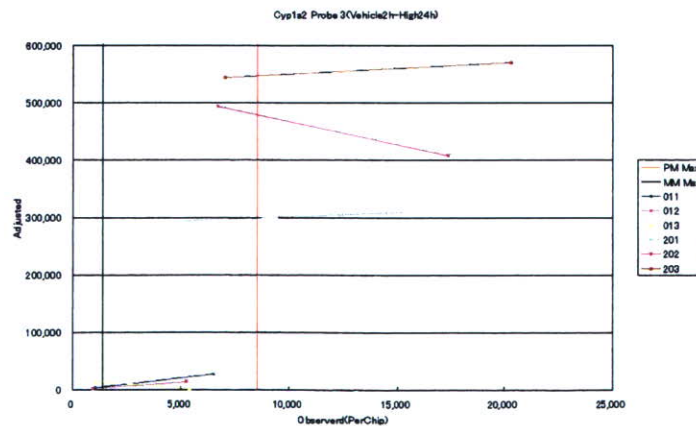
計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008

1.10.6. 全プローブ計算⑥検証②プローブ単位



- Cyp1a2のプローブ単位でチェック
 - 各線が各チップを表す



計測値はLangmuir曲線の最大値を大きく上回っている。

Copyright © NTT COMWARE 2008