

全プローブ計算⑧非線形最適化の結果

- ・ 現在までの非線形モデルよりも非線形計算の安定性がよくない。非線形最適化計算中も数値計算上のエラーで多く停止した。

終了状態	Brain側飽和	Liver側飽和	
勾配ベクトル判定収束	63,035	63,856	} 収束した
移動量判定収束	48,817	33,274	
局所的最小値	37,917	22,926	} 局所的収束
総計	149,769	120,056	

全プローブ計算⑧ 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(○)	PM											総計			
	0	1	2	3	4	5	6	7	8	9	10		11	12	20
0		3862	1943	1009	421	190	73	9	3						7510
1	4828	3833	2341	1335	700	291	95	29	8	1					13462
2	2455	2294	1609	1002	541	245	111	34	10	1					8302
3	1303	1163	868	563	314	156	58	14	6	1			1		4449
4	676	591	416	247	108	55	18	11	6	1					2128
5	352	287	159	85	30	16	4	3	1		1	1			939
6	111	72	47	20	14	4	1								269
7	22	18	16	5	3	2	1								67
8	7	5	1	1	1										15
9	1	2													3
11													1		1
総計	9796	12127	7400	4267	2132	859	361	100	36	4	1	1	1	1	37146

PM側で11プローブが対象となるものがなく、適切な補正は難しい

全プローブ計算⑧ Langmuir Model採用個数

- Liver側飽和

PM

合計 / Count(*)	列ラベル	0	1	2	3	4	5	6	7	8	9	10	11	13	17	総計
行ラベル																
0		4421	1203	350	163	102	49	47	23	5	7	2				6372
1		6462	3344	1216	432	225	136	73	57	44	19	8	5			12021
2		1365	1117	575	267	140	123	121	71	60	34	27	5			3920
3		179	214	166	106	91	92	87	77	55	46	25	12			1149
4		30	38	39	42	44	62	70	85	85	75	45	14			612
5		3	8	12	17	27	38	57	78	76	68	56	28			468
6		2	3	6	5	16	18	35	55	50	64	58	27			351
7		1	2	3	3	8	10	21	35	65	51	64	44			305
8				2	3		7	16	23	31	51	74	88			297
9			1			1	1	5	12	20	29	48	127			245
10							2	2	5	6	15	61	205			296
11									1	2	7	23	228	1		259
13															1	1
総計		8061	9148	3224	1225	715	591	541	546	497	464	491	792	1	1	26297

MM

PM/MM双方で1プローブが補正対象となるものが見つかり、プローブセットとして、飽和現象が存観測されたと考えられる。

全プローブ計算⑧ 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てプロットし、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

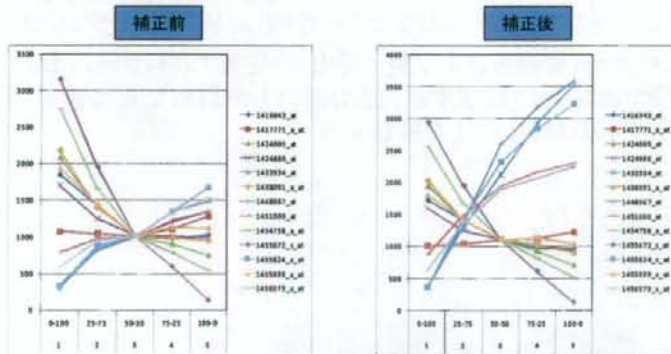
全プローブ計算⑧検証①

- 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

全プローブ計算⑧検証①

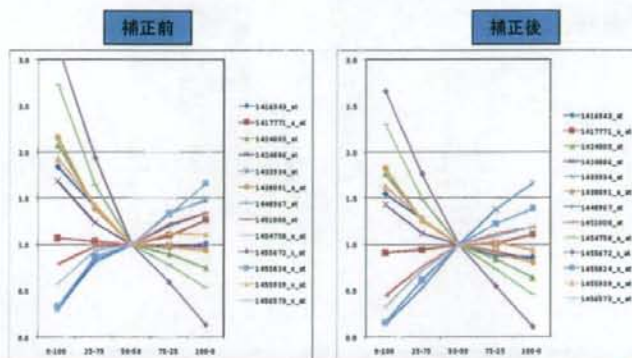
- 対象プローブセット:50%:50%において、1000程度の値を持つ
- 補正前と補正後の比較(Intensity)



直線性の示す補正結果となった。

全プローブ計算⑧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
- 補正前と補正後の比較(50%:50%で正規化)



直線性を示す補正結果となった。

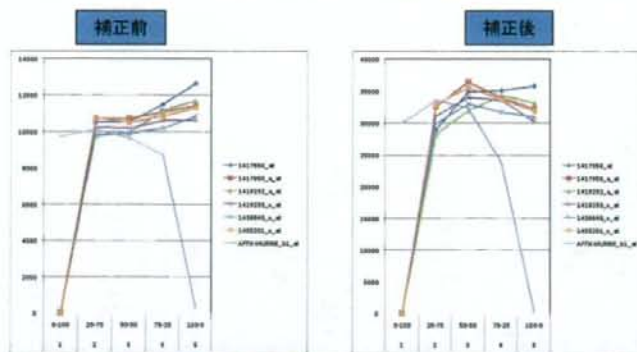
全プローブ計算⑧検証①

- 対象プローブセット: 50%:50%において、10000程度の値を持つ

affy_id	Intensity
1417556_at	10550.2
1417950_at	10735.1
1419232_at	9940.1
1419233_x_at	10193.0
1434110_x_at	10519.8
1438840_x_at	9940.5
1455201_x_at	10504.2
AFFX-MURINE_b1_at	9686.6

全プローブ計算⑧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



補正により、0%-25%-50%は、若干直線に引き延ばすことができ、50%-75%-100%において、大きな変化がなく良好な結果となった

まとめ

Langmuir吸着等温式を用いてある程度飽和に現象に関する補正が可能であった。

- 課題
 - LBMにおいて飽和現象を示したことのあるプローブだけが補正対象となる
 - 当該LBM以外での適用に関して不安定である
 - チップごと補正が不安定である
 - ・ バックグラウンドノイズ、最大輝度
- 計算上必要な項目
 - 補正計算において、飽和値以上を示した場合の対処方法
 - ・ ある一定値以上は線形に補正されるとみなす
 - 非線形収束において、高値補正の項目は考慮に入れない
 - ・ 線形の項目をいれると収束の安定性に欠ける

6. プローブ配列に基づく マイクロアレイデータ補正

Cross-Hybridizationの影響と考えられるの数式モデルの作成

データ補正の必要性

- 化合物間クラスタリング(H18実施)において、高発現のプローブセットの差異によって、クラスタが構成され系統的な誤差が疑われた。
- 各プローブセットで濃度相関を調べたところ飽和と考えられる現象が認められ、Langmuirの吸着等温式を用いた飽和補正を試みた(H19実施)。

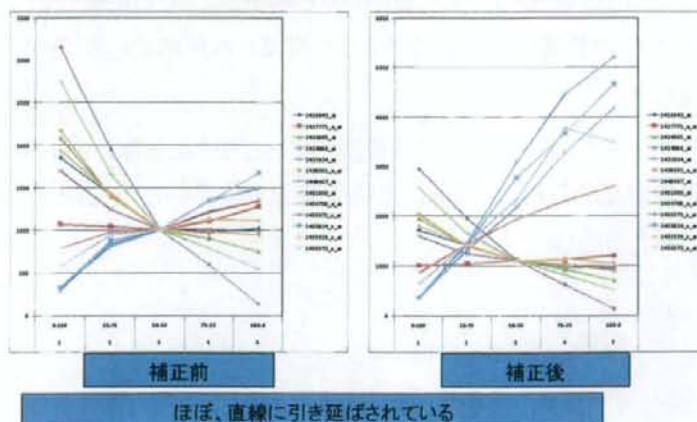
飽和補正(単一Langmuir)

• H19年度の研究を継続し、次の条件を用いた補正を実施した。

1. Langmuirの吸着等温式を補正に用いる。
2. 高濃度域を直線で変換する。
3. MMの補正係数もPMで求める。
4. MASSのサマライズ計算における飽和上限値を撤廃する。

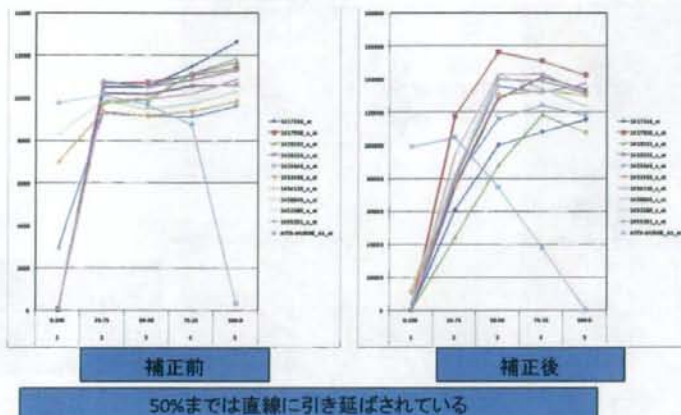
単一Langmuir飽和補正結果(LBM)

• 線形性チェック: 中発現域(50%:50%で1000程度)



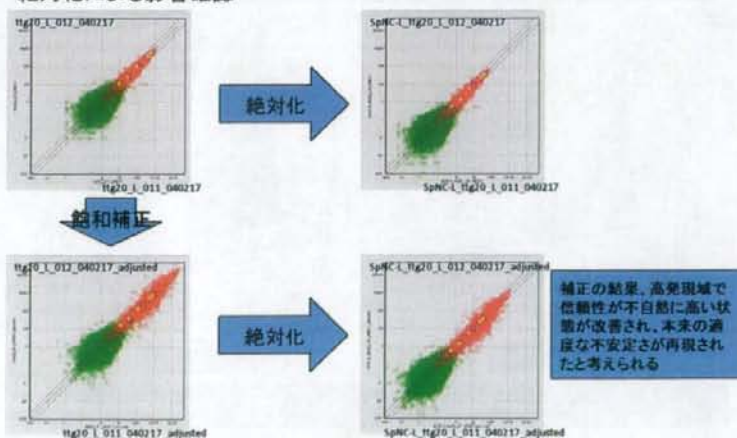
単一Langmuir飽和補正結果(LBM)

- 線形性チェック: 高発現域(50%:50%で、10000以上)



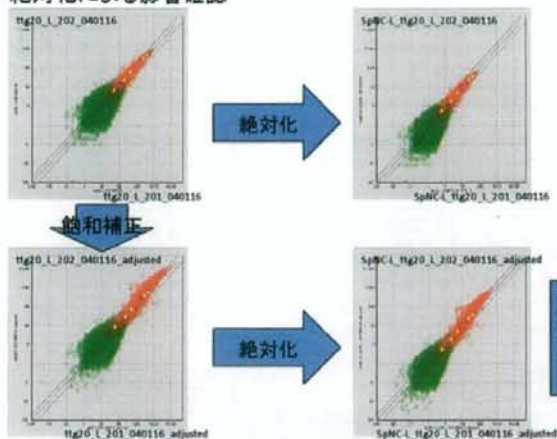
単一Langmuir飽和補正結果 2チップ間散布図(TTG020, 24H-HighDose)

- 絶対化による影響確認



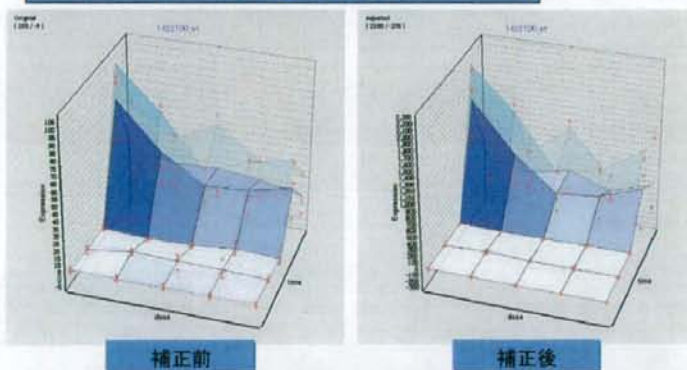
単一Langmuir飽和補正結果 2チップ間散布図(TTG020, 24H-HighDose)

- 絶対化による影響確認



単一Langmuir飽和補正結果： TTG20(Cyp7a1:1422100_at)

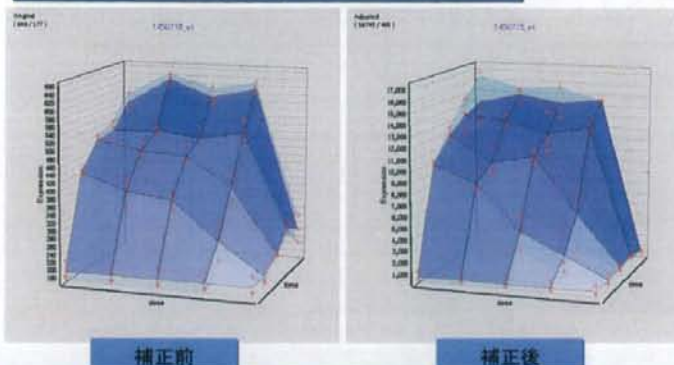
TCDDの影響をうける遺伝子



ほぼ同じ形状であるが、補正計算後非常に大きな値となった。

単一Langmuir飽和補正結果：
TTG20(Cyp1a2:1450715_at)

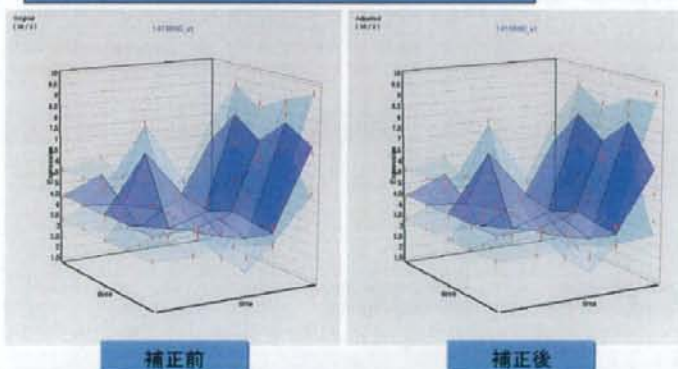
TCDDの影響をうける遺伝子



補正により、24時間での不安定さが増している。また、非常に大きな値となった。

単一Langmuir飽和補正結果：
TTG20(Clock:1418660_at)

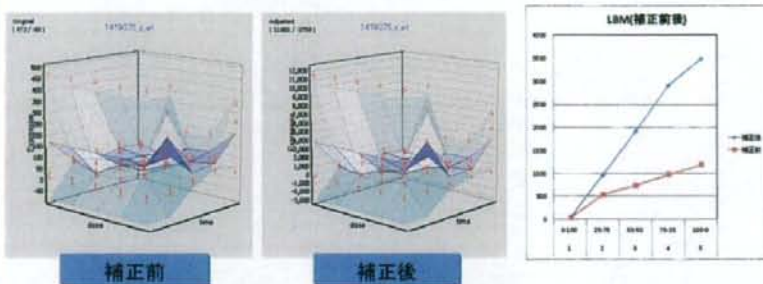
TCDDの影響を受けない遺伝子



発現の小さな遺伝子では、補正による大きな変化はない

単一Langmuir飽和補正結果：
TTG20(Saa1:1419075_s_at)

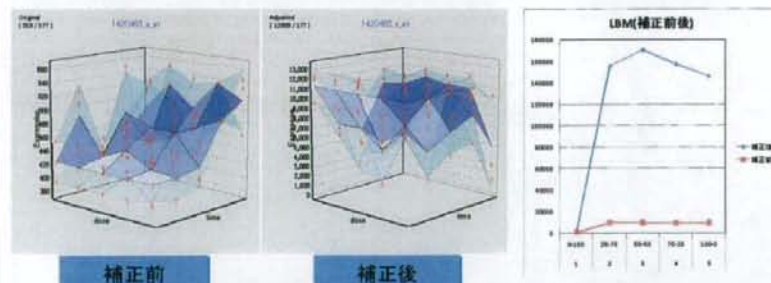
補正後いずれかのチップで20万以上のIntensityを持つ遺伝子



TTG20の結果がLBMの範囲を大きく逸脱しており、外挿計算対象で拡大され、25倍の値になっていると考えられる。

単一Langmuir飽和補正結果：
TTG20(Mup1//Mup2:1420465_s_at)

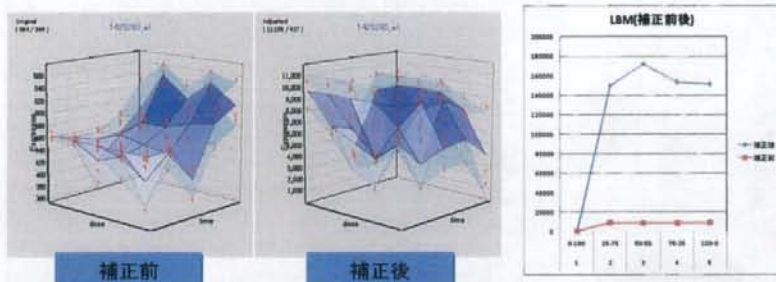
補正後いずれかのチップで20万以上のIntensityを持つ遺伝子



LBMの25%で完全飽和しており、適切な係数の推定ができていないと考えられる

単一Langmuir飽和補正結果:
TTG20(Alb1:1425260_at)

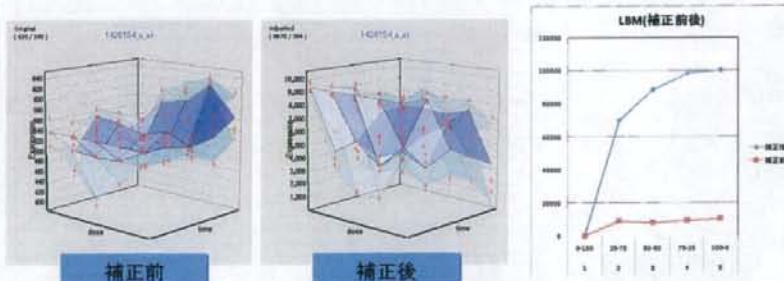
補正後いずれかのチップで20万以上のIntensityを持つ遺伝子



LBMの25%で完全飽和しており、適切な係数の推定ができていないと考えられる

単一Langmuir飽和補正結果:
TTG20(Mup1//Mup3:1426154_s_at)

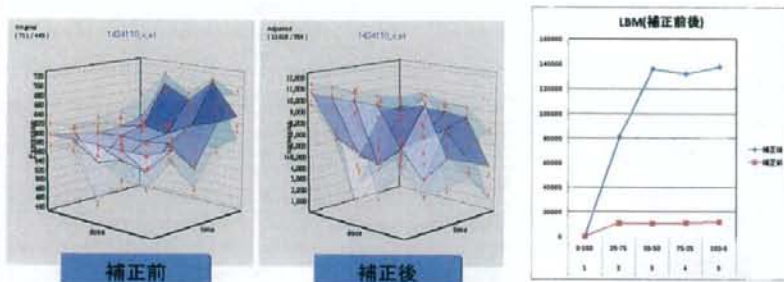
補正後いずれかのチップで20万以上のIntensityを持つ遺伝子



LBMの25%で完全飽和しており、適切な係数の推定ができていないと考えられる

単一Langmuir飽和補正結果：
TTG20(Mup1:1434110_x_at)

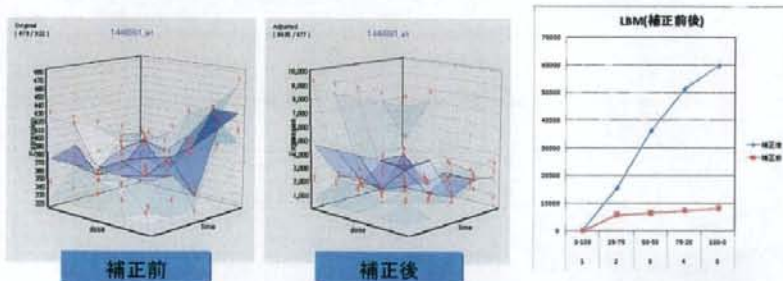
補正後いずれかのチップで20万以上のIntensityを持つ遺伝子



LBMの25%で完全飽和しており、適切な係数の推定ができていないと考えられる

単一Langmuir飽和補正結果：
TTG20(Hp:1448881_at)

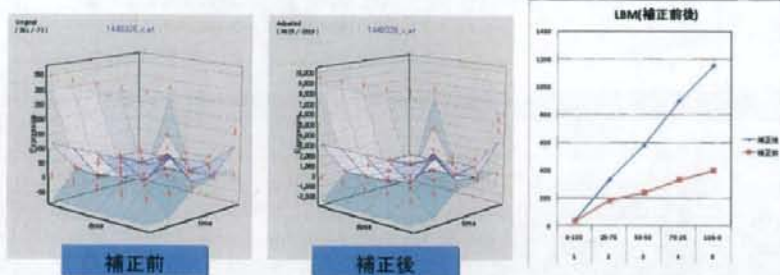
補正後いずれかのチップで20万以上のIntensityを持つ遺伝子



LBMの25%で完全飽和しており、適切な係数の推定ができていないと考えられる

単一Langmuir飽和補正結果： TTG20(Saa2:1449326_x_at)

補正後いずれかのチップで20万以上のIntensityを持つ遺伝子

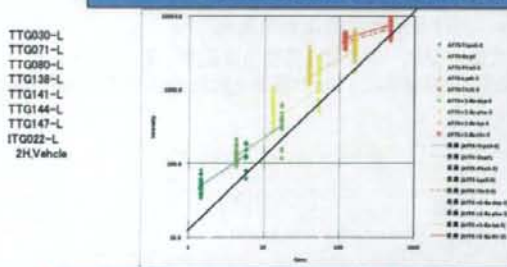


単一のマウスで発生した大きな値なので、補正計算に反映できていないと考えられる

スパイク飽和現象確認

- スパイクも飽和していることから、スパイクも飽和補正すれば、一般プローブの補正結果がより真の値に近づくのではないか？

スパイク濃度と発現値(RAW)の対応関係



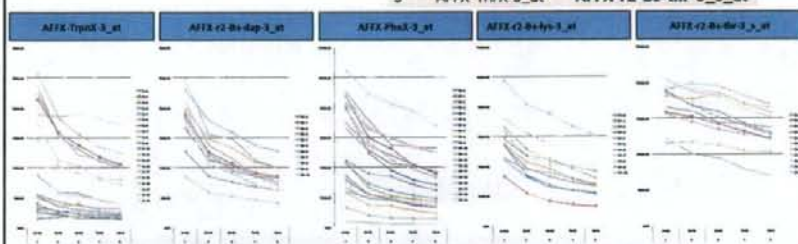
GSC添加量を変更している前後の実験で、化合物投与の影響を受けない2時間溶媒対照群データを用い、GSCのIntensityを比較した。

比例直線より遠れているので、スパイクの各プローブセットは飽和の傾向を示し始めている。しかし、一般プローブの補正結果を数十倍に押し上げる影響があるとは考えられない。

スパイク飽和現象の確認

スパイクデータがLBMでどのように変動しているかを確認した。

	Old Probe Set	New Probe Set
1	AFFX-TrpnX-3_at	-
2	AFFX-DapX-3_at	AFFX-r2-Bs-dap-3_at
3	AFFX-PheX-3_at	AFFX-r2-Bs-phe-3_at
4	AFFX-LysX-3_at	AFFX-r2-Bs-lys-3_at
5	AFFX-ThrX-3_at	AFFX-r2-Bs-thr-3_s_at



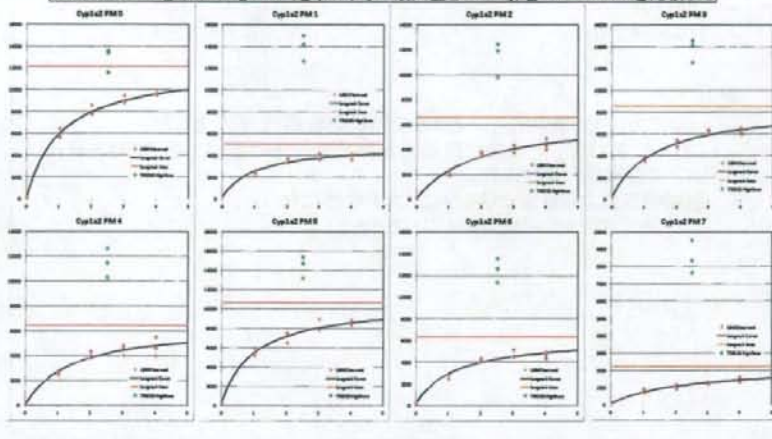
本来、スパイクは変動がないはずであるが、Liver、Brainの特性から添加量が異なり、若干の変動を見せており、最大スパイクで飽和と思われる現象が示唆されている

単一Langmuir補正の不整合

- LBMで飽和しており、TCDDで更に発現量が増加するRNA(Cyp1a2)の挙動を確認した
 - Cyp1a2(1450715_at:10092)のLBMに基づくLangmuir曲線とTTG020-L(TCDD)のHigh-Doseにおける発現値の比較を実施
 - Langmuir曲線の最大値がプローブごとに異なる原因が、ターゲット以外のRNAの吸着によるものであれば、LBMで推測された最大値を超えた値が観測される。

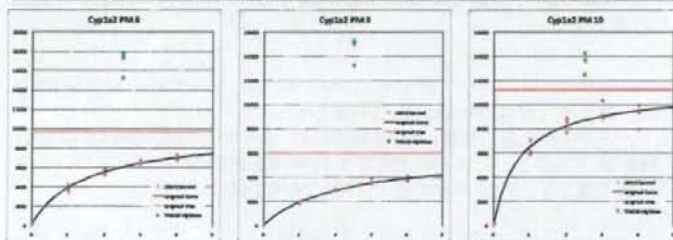
単一Langmuir補正の不整合(結果①)

LBMで飽和しているにもかかわらず、TCDDで更に発現量が増加するRNA(Cyp1a2)の挙動
LBMのrawdataをもとにLagmuir式を求め、その最大値とTCDD影響でのrawdataを比較した



単一Langmuir補正の不整合(結果①)

LBMで飽和しているにもかかわらず、TCDDで更に発現量が増加するRNA(Cyp1a2)の挙動
LBMのrawdataをもとにLagmuir式を求め、その最大値とTCDD影響でのrawdataを比較した



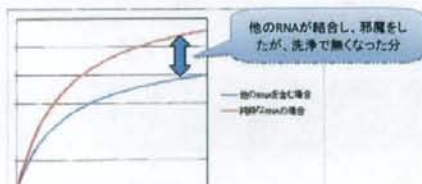
Probe 0以外は、TTG020の最大用量・最大時刻における発現量がLBMによるLangmuir曲線の推定最大値を大きく上回っている。TCDDの誘導により、mRNAの発現量がLBMでカバーされている数値範囲を超えたために、飽和最大値を超えたと考えられる

まとめ(単一Langmuir飽和補正)

- 次の改良を施し、飽和補正計算を実施した。
 - LBM
 - TTG20(TCDD)
- 結果
 - LBMにおいて、飽和しているような形状の直線性が向上した。
 - TTG20において、化合物で誘導のかかる遺伝子においても、形状を維持しつつ直線性が向上した。
 - 飽和補正により想定よりもかなり大きな値になった。
 - 一桁大きなコピー数となるものがあった。

仮説: RNAの競合吸着

- 各RNAに対して、吸着サイトとターゲットRNAの1対1対応を前提としたLangmuir吸着等温方程式で飽和補正を実施すると、補正値が大きめに推定されてるのではないか
- 仮説
 - 一つの吸着サイトに、ターゲット以外のRNAが結合し、ターゲットの結合ができなくなり、見掛け上最大飽和値が小さく推定される。
 - ターゲットRNAが大量に発現している場合には、ターゲット以外のRNAが結合する間もなく本来のターゲットRNAで吸着サイトが埋まってしまうため、見掛け上の最大飽和値がターゲットのみの場合に近くなるのではないだろうか？



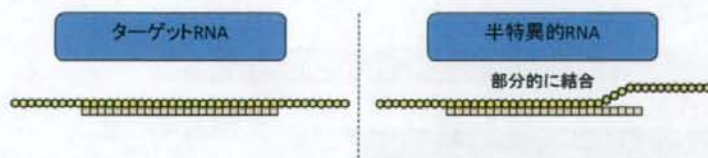
補正係数算出のためのLBMは、何らかの工夫が必要である。

Cross-HybridizationとWashingによる 見かけ上の最大値の変化の考察

- LBMで推定したLangmuir方程式を用いて、実データ(TTG020)の補正を試みたが、推定最大値を超過した遺伝子が存在した。この現象は特に、TTG020で大きく誘導のかかる遺伝子で顕著であった。
- 原因として、次のものが想定できる
 - スパイクの飽和
スパイク自身が飽和し、大きな値を小さくしてしまう
 - Cross-Hybridizationによる疑似的飽和
半特異的結合すれば対象RNAの結合を阻害するが、洗浄工程で流され、阻害分だけ最大値が下がっているように観測される
 - Hybridization時間
最初に半特異的なものも含めて結合し、ゆっくりと特異的なRNAに置き換わっていく。この時間が不十分なのではないか

Cross-HybridizationとWashingによる 見かけ上の最大値の変化の考察

- 想定される状況

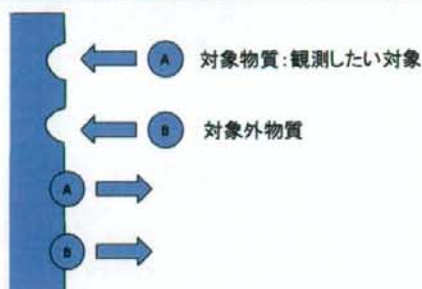


仮説

半特異的RNAは、洗浄工程で落とされ発現値として計測されるため、見掛け上の最大値(飽和値)を引き下げる。また、補正式パラメータ設定用の標準サンプル(LBMサンプル)と比べ、ターゲットRNAが極端に多い場合には、補正式で想定される最大値を大幅に上回る

Langmuir吸着等温式の2溶質への拡張

Langmuirの基本に戻り、2溶質への拡張を行う



吸着速度と離脱速度が平衡状態となると仮定して導き出されたのが、Langmuirの方程式、2溶質で式を組立てて解いてみる。

Langmuir吸着等温式の2溶質への拡張

Langmuir方程式の原理（1溶質の場合）

吸着速度は、吸着体の残り個数と溶質の濃度に比例すると仮定する

$$V' = K \cdot (N - n) \cdot C$$

溶質の離脱速度は、吸着した個数に比例する

$$V = \alpha \cdot n$$

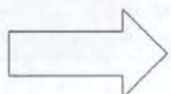
吸着と離脱の速度が平衡していると仮定する

$$\alpha \cdot n = K \cdot (N - n) \cdot C$$

$$\alpha \cdot n + K \cdot n \cdot C = K \cdot N \cdot C$$

$$n \cdot (\alpha + K \cdot C) = K \cdot N \cdot C$$

$$n = \frac{K \cdot N \cdot C}{\alpha + K \cdot C} = N \cdot \frac{\frac{K}{\alpha} \cdot C}{1 + \frac{K}{\alpha} \cdot C}$$



$$n = N \cdot \frac{k \cdot C}{1 + k \cdot C}$$

表記

- N : 吸着体個数
- n : 溶質が吸着している個数
- C : 溶質の濃度
- α : 吸着速度比例定数
- K : 離脱速度比例定数

Langmuir吸着等温式の2溶質への拡張

2溶質の場合

吸着速度は、吸着体の残り個数と溶質の濃度に比例すると仮定する

$$V' = K \cdot (N - n_A - n_B) (C_A + C_B)$$

個々の溶質の吸着速度は、溶質の濃度の比に比例する

$$V'_A = \frac{C_A}{C_A + C_B} V'$$

$$V'_B = \frac{C_B}{C_A + C_B} V'$$

個々の溶質の離脱速度は、吸着した個数に比例する

$$V_A = \alpha \cdot n_A$$

$$V_B = \beta \cdot n_B$$

Langmuir吸着等温式の2溶質への拡張

2溶質の場合（式展開）

平衡を仮定する

$$\begin{cases} K \cdot (N - n_A - n_B) C_A = \alpha \cdot n_A \\ K \cdot (N - n_A - n_B) C_B = \beta \cdot n_B \end{cases}$$

式を展開すると

$$\begin{cases} K \cdot (N - n_B) C_A = (\alpha + K \cdot C_A) n_A \\ K \cdot (N - n_A) C_B = (\beta + K \cdot C_B) n_B \end{cases}$$

$$\begin{cases} n_A = \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_B = \frac{K \cdot (N - n_A) C_B}{(\beta + K \cdot C_B)} \end{cases}$$

濃度と吸着量の関係式

片方の吸着量を消すことが可能である

溶質BがWashingで流され観測不能であると仮定し、観測不能なnBを消去する

$$\begin{aligned} n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \\ n_A &= \frac{K \cdot (N - n_B) C_A}{(\alpha + K \cdot C_A)} \end{aligned}$$