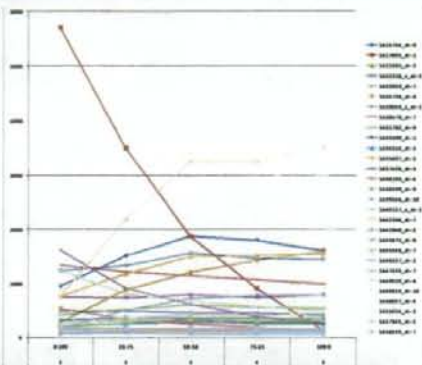


## 2.2.8.飽和しているRNAと似ている配列を有するProbeの挙動



Alb1と似ている配列: Alb1と5末端で12mer以上連続して一致するProbe

Probe ID	Score	Rank	Category
131	85	5	14, 12
232	1306	1	14, 12
229	4000	3	14, 12
254	6022	2	14, 12
8	8259	7	12, 14
414	2032	8	14, 12
17	2566	0	14, 12
34	3395	1	1, 25
64	3566	3	1, 25
94	3566	4	1, 25
114	3566	2	1, 25
246	3566	6	1, 25
214	3566	7	1, 25
260	3566	5	1, 25
362	3566	9	1, 25
453	3566	10	1, 25
377	12353	1	13, 13
56	12394	2	14, 12
136	16046	3	14, 12
130	17515	1	14, 12
327	18023	2	14, 12
48	19113	2	14, 12
131	21232	3	14, 12
202	22899	4	14, 12
183	22645	3	14, 12
217	23821	4	14, 12
461	24833	1	14, 12
31	26362	2	14, 12
362	27254	2	14, 12
316	27978	3	13, 13
8	27983	2	14, 12
252	27983	2	14, 12
302	31741	2	14, 12
67	33319	6	12, 14
23	34248	10	13, 13
349	34202	4	14, 12
4	35979	4	14, 12
31	42242	3	14, 12
217	42728	7	12, 13



Brain側で飽和しているProbeは凹な形状を示し、Liver側で飽和しているProbeは飽和最大値が低い

Copyright©2008-2009 NTT DATA Corporation

## 2.2.9.高発現遺伝子の影響を受けていないと考えられるプローブ



- 高発現遺伝子の影響を受けていないと考えられるプローブは、LBMに対して、直線性を維持しているはずである。このようなプローブを見つけ出し、直線性を確認する

LBMにおいて、3マウス(1実験条件)の発現平均が閾値以上のプローブ

閾値以上のプローブ→プローブセット(高発現遺伝子とみなす)

高発現遺伝子(ターゲット配列)と8個以上連続で合致するプローブ(影響を受けているプローブ)

全プローブから、影響を受けているプローブを取り除き、影響を受けていないプローブを見つける

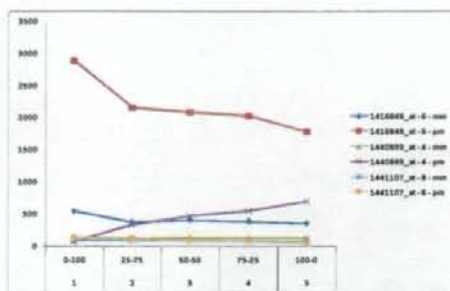
閾値	影響を受けていないプローブ数
12000	0
14000	3
16000	152
18000	2034

Copyright©2008-2009 NTT DATA Corporation

## 2.2.9.高発現遺伝子の影響を受けていないと考えられるプローブ



- 14000以上の発現値を有するターゲットと、8mer以上連続で合致しないプローブ - 全3プローブ



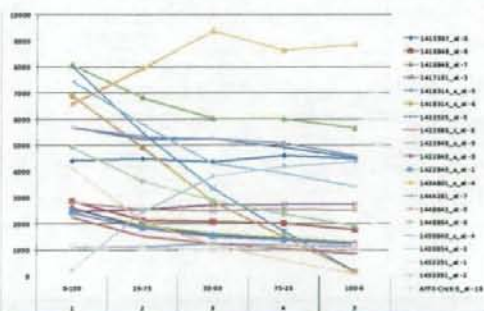
発現量が低く、直線性を維持しているかの、判断が難しい

Copyright(C)2009-2009 NTT DATA Corporation

## 2.2.9.高発現遺伝子の影響を受けていないと考えられるプローブ



- 16000以上の発現値を有するターゲットと、8mer以上連続で合致が存在しないプローブ - 平均値が1000以上のプローブ(PM)



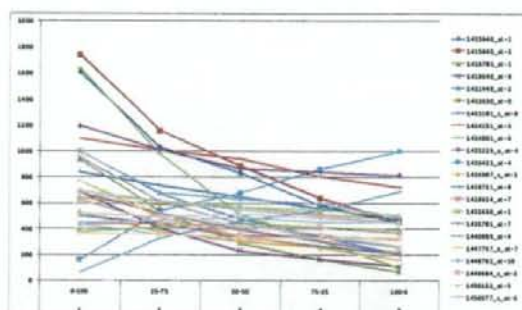
Liver側で飽和しているプローブも存在するが、概ね直線と考えられる。

Copyright(C)2009-2009 NTT DATA Corporation

## 2.2.9.高発現遺伝子の影響を受けていないと考えられるプローブ



- ・ 16000以上の発現値を有するターゲットと、8mer以上連続で合致しないプローブ
  - 平均値が300-1000のプローブ (PM)



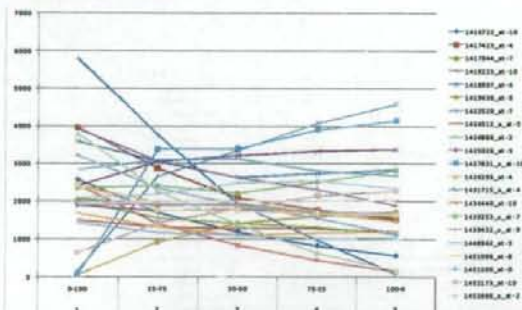
Liver側で飽和しているプローブも存在するが、概ね直線と考えられる。

Copyright©2008-2009 NTT DATA Corporation

## 2.2.9.高発現遺伝子の影響を受けていないと考えられるプローブ



- ・ 対比用: ランダムに選んだ200プローブ
  - 平均値が1000-3000のプローブ (PM)



ランダムに選択されたプローブでは、最大値が2000以下で飽和していると考えられる形状を示すものもあり、8mer連続合致を避けることで、直線性の高いプローブが導き出されたと考えられる。

Copyright©2008-2009 NTT DATA Corporation

## 2.2.10.Affymetrix配列設計の網羅的検証

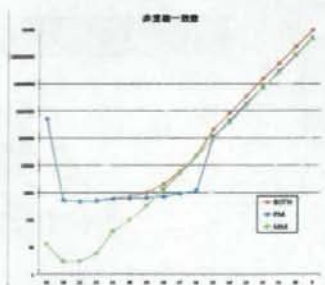
- ・ Affymetrix社が設計した配列に対して、部分的な配列の一致に関して網羅的なチェックを行い、どの程度Cross-Hybridizationが発生しているようなのかをチェックする。
- ・ 対象
  - GeneChip MOE430-2
  - プローブ配列(PM) vs. ターゲット配列
  - 連続9mer以上の一致を検出
    - ・ 断続的な一致は、阻害要因として可能性が低いと考え、9mer以上に設定

## 2.2.10.Affymetrix配列設計の網羅的検証 ターゲットとプローブ間の一致数

ターゲットとプローブの間で、連続して一致する状況の個数を調査した。

MMに25mer一致が13個存在した。

連続数	PM	MM	PM/MM合計
25	526,104	13	526,117
24	527	3	530
23	481	3	484
22	513	6	519
21	592	40	632
20	631	102	733
19	668	331	999
18	730	1,338	2,068
17	948	5,143	6,091
16	1,199	21,996	23,195
15	117,190	92,201	209,391
14	440,156	384,475	824,631
13	1,774,096	1,589,142	3,363,238
12	7,217,804	7,737,802	14,955,606
11	29,355,970	27,443,556	56,799,526
10	118,411,755	112,115,989	230,527,744
9	477,288,011	455,736,406	933,024,417



12mer連続一致では、MM側が若干多い。MMの設計で中央位置を除くというのが影響していると考えられる

16mer連続一致では、PM側が圧倒的に少ない。設計段階でチェックされていると思われる。

## 2.2.10.Affymetrix配列設計の網羅的検証 ターゲットとプローブ間的一致数(端)



プローブの端は、その先で反発力を生み出すものがないので、中央の場合よりも、Hybridization時点において、離脱速度が低く、大きな阻害要因となると考えられる。状況を確認する

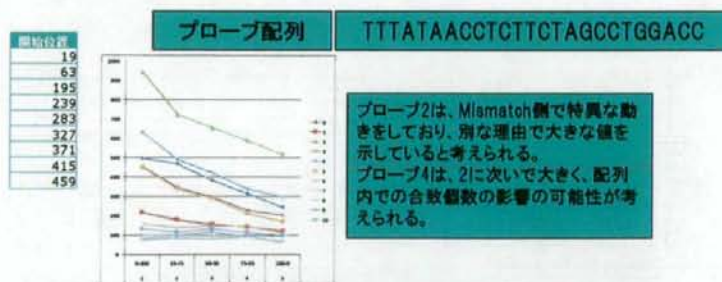
プローブ長	一致数	中央	両末端部	両端部
25	526,117	0	526,117	526,117
24	530	0	281	249
23	484	10	252	222
22	519	34	235	250
21	632	55	227	350
20	733	142	230	361
19	999	270	351	378
18	2,068	1,017	512	539
17	6,091	3,900	1,032	1,159
16	23,195	16,430	3,282	3,483
15	209,391	147,445	29,890	32,056
14	824,631	625,689	94,639	104,303
13	3,363,228	2,631,294	354,237	377,697
12	14,955,606	11,038,939	1,906,381	2,010,286
11	56,799,526	46,111,992	5,180,454	5,507,080
10	230,527,744	190,471,482	19,355,046	20,701,216
9	933,024,417	782,011,343	73,304,200	77,708,874

Copyright©2008-2009 NTT DATA Corporation

## 2.2.10.Affymetrix配列設計の網羅的検証 特異なプローブ設計例 ①



- ターゲット上に複数合致
- 1455452\_x\_at(AI449310)のプローブ4は、ターゲット配列で9か所合致する。



25merで特異的に結びつくことを前提にしているのので、このようなプローブは設計すべきではないし、必要な場合にはその影響を評価しサマライズ処理で反映すべきである

Copyright©2008-2009 NTT DATA Corporation



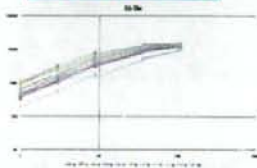
## 2.2.11.Percellome用GSCの飽和



GSC(Graded Spike Cocktail)の飽和程度を確認するために、GSCの量を変化させ、実験を行った。

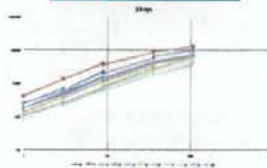
- 飽和現象は、最上位(Thr)、2番目(Lys)では明らかに見られた。また、3番目(Phe)でも一部観測された。

最高濃度(Thr)



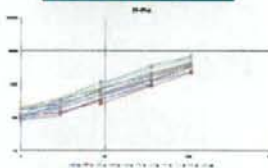
旧濃度で飽和していることが見ただけで判別可能である

第2濃度(Lys)



旧濃度で若干の飽和現象と思われる現象が発生している

第3濃度(Phe)



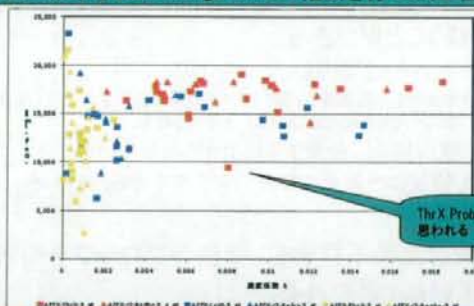
低用量で少し大きめの値が観測されている

Copyright©2006-2009 NTT DATA Corporation

## 2.2.11.Percellome用GSCの飽和



GSC飽和実験の結果をもとにLangmuir式の推測を行った(上位3濃度)



多溶質のLangmuir方程式から導き出される結果と合致した

1. 変化しない背景がある場合には、最大濃度への影響はなく、濃度が薄められた曲線を描く。→ Thrは、この条件に合致するとみなせる

2. 運動して変化する溶質が存在する場合には、最大量が抑えられる。

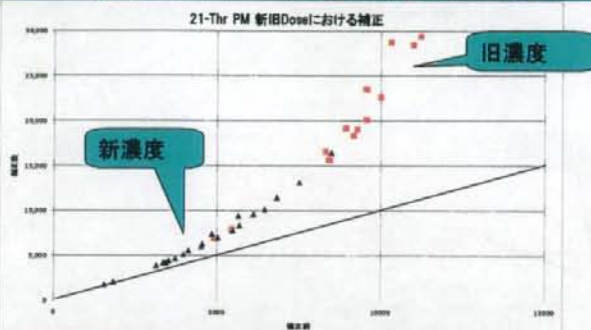
→ LysやPheは、Thrの影響を受けており、この条件に合致するとみなせる

Copyright©2006-2009 NTT DATA Corporation

## 2.2.11.Percellome用GSCの飽和

### 最高濃度Thrの飽和程度

各ブロープのLangmuir式の推定を行い、新旧濃度における飽和補正を実施した



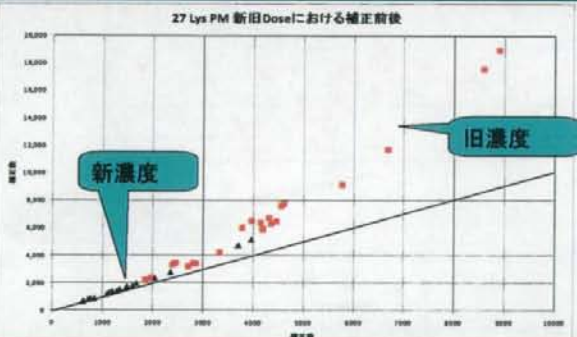
旧濃度では、補正後2倍程度になっていたが、新濃度では、50%増程度に収まっている

Copyright©2009-2009 NTT DATA Corporation

## 2.2.11.Percellome用GSCの飽和

### 2番目Lysの飽和程度

各ブロープのLangmuir式の推定を行い、新旧濃度における飽和補正を実施した



旧濃度では、補正後50%増程度になっていたが、新濃度では、5%増程度に収まっている

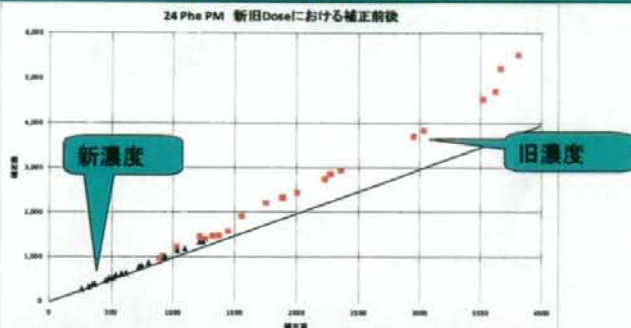
Copyright©2009-2009 NTT DATA Corporation



## 2.2.11. Percellome用GSCの飽和

### 3番目Phe飽和程度

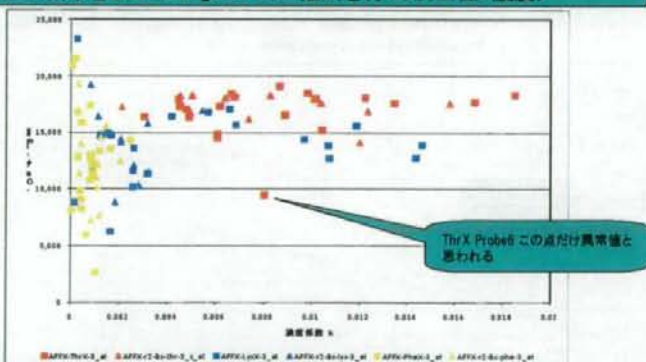
各プローブのLangmuir式の推定を行い、新旧濃度における飽和補正を実施した



旧濃度では、直線近似や補正後10%増程度になっていたが、新濃度では、若干の増加(1%増程度)に収まっている

## 2.2.12. Percellome用GSCの飽和実験 (新旧GSCプローブ)

GSC飽和実験の結果をもとにLangmuir式の推測を行った(上位3濃度)



最大濃度(thr)は、新旧プローブともほぼ同様な最大値を示している。

第2(Lys)、第3(Phe)の濃度は、若干小さな最大値を示している。

濃度係数が小さい場合は、実験上、高い飽和濃度までいってならず、高濃度の測定が大きいことを示している

## 2.2.12.Percellome用GSCの飽和実験 (新IGSCプローブ)



ThrXのプローブ6のImax低下の原因を探る

Thrの旧プローブセットの第6プローブだけが、Imaxが小さい。何らかの原因があると考えられる。現在の仮説では、GSCが25未満で結合しているため、洗い流されていることに起因していると考えている。そのような状況が見つかれば、仮説の補強材料として利用可能であり、25未満の結合力の評価として利用可能である。

ThrXの各プローブで、連続一致するターゲットの個数を示す。

プローブ番号	8	9	10	11	12	13	14	15	17	25
0	3133	668	119	38	3	2	1			1
1	1914	387	52	8	1					1
2	2081	471	98	20	5		2	1		1
3	3183	806	211	59	8	4	2			2
4	1219	274	36	12	1					2
5	2444	627	172	26	9	2	1			2
6	5570	1621	421	118	30	12	3	2		2
7	3245	770	181	54	5	3				2
8	1360	256	45	13	2					2
9	2667	733	177	53	9	3	1			2
10	2705	634	141	35	14	3				2
11	3587	948	203	55	17	3				2
12	2705	622	126	41	6	2				2
13	2407	507	138	27	7		1			2
14	1605	347	59	14	2					2
15	2804	636	168	24	11	3				2
16	1890	450	61	15	1	1				2
17	2250	456	99	14	6	2				2
18	2361	524	133	30	10	2				2
19	3612	928	224	64	7	6	2		1	1

第6プローブは、他のプローブよりも多く一致している。  
仮説では、検体用プローブに対する一致では、Imaxが減少しないが、状況把握のため、確認した。

Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.12.Percellome用GSCの飽和実験 (新IGSCプローブ)



ThrXのプローブ6のImax低下の原因を探る

ThrXの全プローブと8mer以上連続で一致したGSCターゲットの全リストを示す

プローブ番号	開始位置	長さ	終了位置	ターゲットID	ターゲット長さ	ターゲット名	備考
19	1	17	17	43081	393	APFX-v2-8a-thr-3_s_at	
19	11	9	19	43081	1	APFX-v2-8a-thr-3_s_at	
19	9	9	17	43080	364	APFX-v2-8a-thr-5_at	
19	17	8	24	43094	516	APFX-ThrX-5_at	
19	17	8	24	43082	514	APFX-v2-8a-thr-5_s_at	
19	7	9	14	43092	138	APFX-ThrX-3_at	
19	8	9	17	43081	30	APFX-v2-8a-thr-3_s_at	
19	8	9	17	43093	252	APFX-ThrX-3_at	
19	13	8	21	43094	634	APFX-ThrX-5_at	
19	13	8	21	43081	186	APFX-v2-8a-thr-3_s_at	
19	13	8	21	43081	51	APFX-v2-8a-thr-3_s_at	
7	4	8	11	43095	132	APFX-ThrX-4_at	
7	6	8	13	43093	97	APFX-ThrX-3_at	
7	3	8	10	43093	202	APFX-ThrX-3_at	
7	4	8	11	43083	82	APFX-v2-8a-thr-4_s_at	
7	4	8	10	43081	156	APFX-v2-8a-thr-3_s_at	
8	17	8	24	43094	582	APFX-ThrX-5_at	
17	17	8	24	43082	582	APFX-v2-8a-thr-5_s_at	
19	5	8	12	43066	101	APFX-ThrX-5_at	
19	5	8	12	43079	13	APFX-v2-8a-thr-5_at	
11	16	8	23	43093	455	APFX-ThrX-3_at	
17	1	8	8	43093	455	APFX-ThrX-3_at	
14	1	8	8	43094	379	APFX-ThrX-5_at	
14	1	8	8	43082	377	APFX-v2-8a-thr-5_s_at	
10	14	8	21	43082	325	APFX-v2-8a-thr-5_s_at	
10	14	8	21	43094	327	APFX-ThrX-5_at	
10	7	8	14	43065	184	APFX-ThrX-3_at	
10	7	8	14	43078	172	APFX-v2-8a-thr-3_s_at	
10	17	8	24	43093	222	APFX-ThrX-3_at	
10	17	8	24	43081	174	APFX-v2-8a-thr-3_s_at	

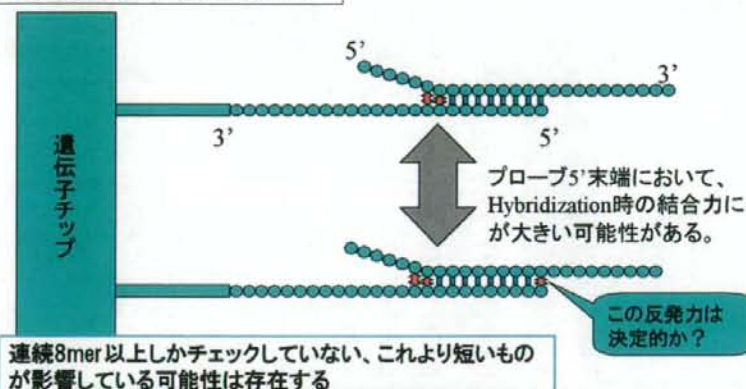
プローブ6だけに存在する特徴として、終了位置25が存在する。つまり、プローブの5'末端側でプローブ間の反発を受けにくい可能性が考えられる。

Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.12.Percellome用GSCの飽和実験 (新旧GSCプローブ)



プローブ末端の結合の有無による違い



Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.12.Percellome用GSCの飽和実験 (新旧GSCプローブ)



プローブセット内の発現値差異の原因(Thr)

発現値の差異は存在するが、lmaxの差異がほとんど存在しない。



Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.12. Percellome用GSCの飽和実験 (新旧GSCプローブ)



### HybridizationとWashingにおける結合の数

GSC飽和実験の最大温度(Thr)の結果

Imaxは同一プローブセット内で、  
同じような値を示す。



Washing工程における「結合力」  
は差異がほとんどない。

目標RNAの濃度を極端に高くすれば、Hybridization後に目標のRNAで占められ、吸着したRNA数は同一と考えられる。これらのRNAがWashing工程で、洗い流されても、似たような値になる。ゆえにWashing工程における「結合力」に差異はない。

### Washing工程において

「結合力」は差異がほとんど  
ない。

仮説1

WashingとHybridizationでは、結合力の値が違う

仮説2

Washingでは、差異が現れるほど時間が長い。平衡状態になっていない。

両仮説とも、妥当と思われる

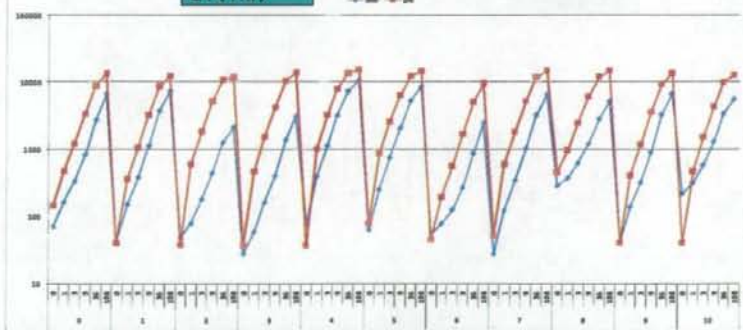
Copyright©2008-2009 JITT, DATA Corporation

## 2.2.13. Percellome用GSCの飽和実験 (PM-MM関係考察)



- PM-MMの関係を確認し、Wash工程における洗浄される程度を考察する

21(Thr)



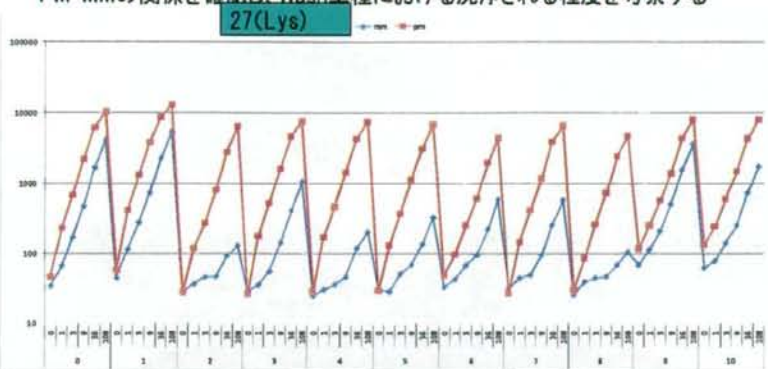
高濃度ではPMが飽和し、低濃度では逆転等がみられるが、MMは、PMの1/10~1/3程度の値で推移している。他の変動成分の影響が少ないため、これが本来の動きに近いと考えられる。

Copyright©2008-2009 JITT, DATA Corporation

## 2.2.13.Percellome用GSCの飽和実験 (PM-MM関係考察)



- PM-MMの関係を確認し、Wash工程における洗浄される程度を考察する



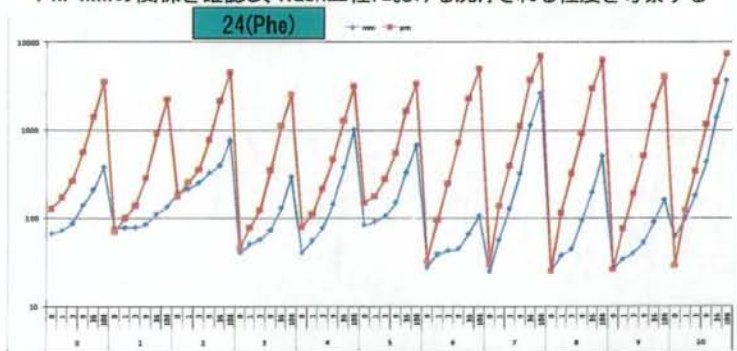
高濃度ではPMが飽和し、低濃度では逆転等がみられ、また、いくつかの例外ブロープが存在するが、MMは、PMの1/10～1/3程度の値で推移している。

Copyright©2008-2009 NTT DATA Corporation

## 2.2.13.Percellome用GSCの飽和実験 (PM-MM関係考察)



- PM-MMの関係を確認し、Wash工程における洗浄される程度を考察する



低濃度で背景と考えられる数値が出ている。また、いくつかの例外ブロープが存在するが、MMは、PMの1/10～1/3程度の値で推移している。

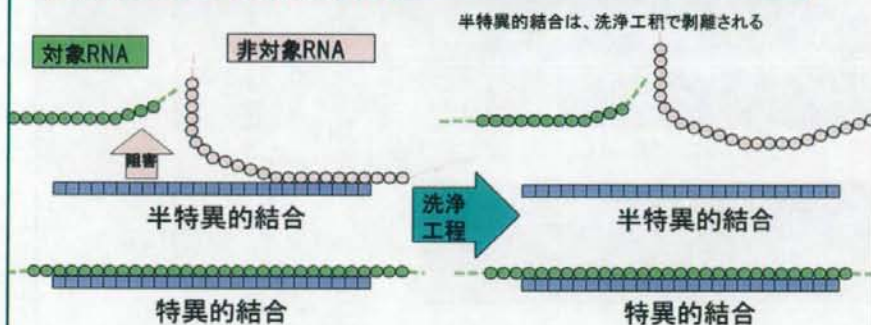
Copyright©2008-2009 NTT DATA Corporation

## 2.2.14.Langmuir吸着等温式の多溶質への拡張



半特異的結合 (Semi-specific Hybridization) という概念を導入し、モデルを構築する

半特異的結合は、プローブと間にある程度の大きさの結合力が働き外れにくく、Hybridization時間内では、特異的結合に置き換えることができない

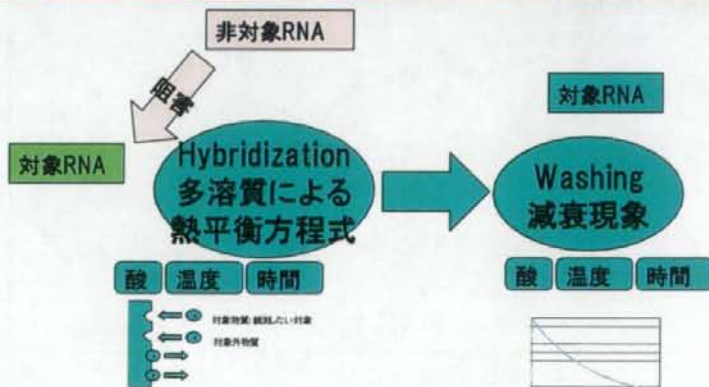


Copyright(C)2008-2009 MTT DATA Corporation

## 2.2.14.Langmuir吸着等温式の多溶質への拡張



各RNAの真の濃度がわかっている場合の計測値を導き出すモデルを仮定する



Copyright(C)2008-2009 MTT DATA Corporation

## 2.2.14.Langmuir吸着等温式の多溶質への拡張



多溶質による吸着のモデル化、Langmuir吸着等温式と同じ仮定

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

$$V' = K \cdot \left( N - \sum_{i=1}^n N_i \right) \left( \sum_{j=1}^n C_j \right)$$

個々の溶質の吸着速度は、溶質の濃度の比に比例する

$$V'_j = \frac{C_j}{\sum_{j=1}^n C_j} V'$$

個々の溶質の離脱速度は、吸着した個数に比例する

$$V_i = \alpha_i \cdot N_i$$

式変形

プローブに吸着したターゲットjの個数

$$N_j = \frac{C_j}{\alpha_j} \cdot K \cdot \left( N - \sum_{i=1}^n N_i \right) + \varepsilon_j$$

プローブに吸着したターゲットjの離脱速度係数  
吸着速度係数は1とし、 $\alpha$ で調整する

$$\alpha_j \geq 0$$

Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.14.Langmuir吸着等温式の多溶質への拡張



計算式の展開  
計算可能な陽形式に変換する

$$N_j = \frac{C_j}{\alpha_j} \cdot K \cdot \left( N - N_j - \sum_{k=1}^m N_{ik} \right) + \varepsilon_j$$

$$N_j + \frac{C_j}{\alpha_j} K N_j = \frac{C_j}{\alpha_j} \cdot K \cdot \left( N - \sum_{k=1}^m N_{ik} \right) + \varepsilon_j$$

$$N_j \left( 1 + \frac{C_j}{\alpha_j} \right) = \frac{C_j}{\alpha_j} \cdot \left( N - \sum_{k=1}^m N_{ik} \right) + \varepsilon_j$$

$$N_j = \frac{N - \sum_{k=1}^m N_{ik}}{\alpha_j + C_j} C_j + \varepsilon_j$$

$$N_j = \frac{N + \hat{N}_j - \sum_{k=1}^m \hat{N}_{ik}}{\alpha_j + C_j} C_j + \varepsilon_j$$

Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.14. Langmuir吸着等温式の多溶質への拡張



各RNAの真の濃度がわかっている場合の計測値を導き出すモデル(数式表現)

①Hybridizationで、プローブ*i*に吸着したターゲット*j*の個数

$$N_{ij} = \frac{N - \sum_{k=1}^m N_{ik}}{\alpha_{ij} + C_j} C_j$$

Washing

②Washing後で、プローブ*i*に吸着したターゲット*j*の個数

$$n_{ij} = N_{ij} \cdot e^{-\gamma \cdot \alpha_{ij}}$$

観測

③Washing後で、プローブ*i*で観測される値

$$n_i = \sum_{j=1}^m n_{ij} + \bar{\varepsilon}_i$$

Copyright©2008-2009 RTT DATA Corporation

## 2.2.15. 離脱速度比例定数（結合強度）の算出



- ・ プローブとRNAの配列から離脱速度を算出できるとすれば、真の濃度を観測値から求めることができる。
- ・ 配列から離脱速度が算出可能か試みた。

Copyright©2008-2009 RTT DATA Corporation



## 2.2.15.離脱速度比例定数（結合強度）の算出



- ・ プローブとRNAの配列から離脱速度の計算方法を、次のように仮定する

$\alpha_i$ : プローブ  $i$  に対する、RNA $_j$  の離脱速度比例定数

小さい値ほど似ている

離脱速度は、結合力の逆数であると仮定する

$$\alpha_i = 1/c_{ij}$$

$c_{ij}$ : プローブ  $i$  に対する、RNA $_j$  の結合強度

結合は連続した配列の長さにより規定可能である。  
長いほど強固である

複数箇所一致する場合は、足し算可能である

$$c_{ij} = f(l_{ij})$$

$l_{ij}$ : プローブとRNA $_j$ の一致する配列の最大長

Copyright © 2008-2009 NTT DATA Corporation

## 2.2.15.離脱速度比例定数（結合強度）の算出



濃度がターゲットと一次の関係で運動していると仮定する

$$C_j = p_{ij} \cdot C_i + q_j$$

1溶質の式

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

lmax: 1溶質 → 多溶質

$$N \longrightarrow N \cdot \frac{1}{\alpha_i p_i + 1}$$

代入する

$$\sum_{j=1}^n \frac{p_{ij}}{\alpha_j} = p_i$$

$$\sum_{j=1}^n \frac{q_j}{\alpha_j} = q$$

$$n_i = N \cdot \frac{1}{\alpha_i p_i + 1} \cdot \frac{c_i}{\frac{K}{\alpha_i p_i + 1} + c_i}$$

最大値は、運動している成分の影響のみを受ける。

Langmuir飽和式は、一定量の成分(背景)の影響を受けるが最大値は影響を受けない。

通常の実験では、きれいなLangmuir飽和式(シグモイド曲線)が描かれ、それ以外の状況を見ることができない。

1. Latin Square (Affymetrix公開データ)
2. Liver Brain Mixture
3. Graded Spike Cocktails 飽和実験

GSC飽和実験の最大濃度RNAの最大値は理論通り、ほぼ一定している。第2濃度の最大値を用いて、離脱速度の検討が可能ではないだろうか

Copyright © 2008-2009 NTT DATA Corporation

## 2.2.15.離脱速度比例定数（結合強度）の算出



GSC飽和実験の第1濃度(Thr)、第2濃度(Lys)に着目し、推定された最大値は、理論値+誤差で示されると仮定する

$$\bar{N} = N \cdot \frac{1}{\alpha_1 p_1 + 1} + \varepsilon$$

$$\bar{N} = N \cdot \frac{1}{\alpha_1 \sum_{j=1}^n \frac{p_j}{\alpha_j} + 1} + \varepsilon$$

$$\bar{N} \left( \alpha_1 \sum_{j=1}^n \frac{p_j}{\alpha_j} + 1 \right) - N = \varepsilon$$

- 第1濃度をターゲットとしてみた場合
    - 第1濃度でターゲット箇所(25mer一致)の場合
    - 第1濃度でターゲット箇所以外(24mer以下一致)の場合
    - 第2濃度の場合
  - 第2濃度をターゲットとしてみた場合
    - 第2濃度でターゲット箇所(25mer一致)の場合
    - 第2濃度でターゲット箇所以外(24mer以下一致)の場合
    - 第1濃度の場合
- 第3濃度以下は無視する

$\alpha$ を配列から計算を行ってみて、誤差  $\varepsilon$  が小さくなりそうな計算式を探る

Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.15.離脱速度比例定数（結合強度）の算出



既知の情報を引数とする結合力に関する関数を組み立てる

Hybridization時点とWashing時点の結合力を求める

次の例を考える

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25				
例1	C	T	T	C	T	T	A	A	C	G	G	T	A	C	T	T	A	T	T	A	C	A	G	T	A	C	G	A	A
例2	A	T	A	T	G	G	T	C	G	A	C	C	C	C	T	G	G	T	T	C	C	A	G	T	A	C	C	A	C
例3	C	C	A	T	G	G	T	C	G	A	T	T	A	G	C	C	G	T	T	A	C	A	G	T	A	C	C	C	A
Probe PM		A	T	G	G	T	C	G	A	T	T	A	G	C	C	G	T	T	A	C	A	G	T	A	C	C			
例1	C	T	T	C	T	T	A	A	C	G	G	T	A	C	T	T	A	T	T	A	C	A	G	T	A	C	G	A	A
例2	A	T	A	T	G	G	T	C	G	A	C	C	C	C	T	G	G	T	T	C	C	A	G	T	A	C	C	A	C
例3	C	C	A	T	G	G	T	C	G	A	T	T	A	G	C	C	G	T	T	A	C	A	G	T	A	C	C	C	A
Probe MM		A	T	G	G	T	C	G	A	T	T	A	G	C	C	G	T	T	A	C	A	G	T	A	C	C			

Copyright(C)2008-2009 NTT DATA Corporation

## 2.2.15. 離脱速度比例定数（結合強度）の算出



プローブとターゲットとの結合力は各塩基での結合力の和により表現可能であると仮定する

各塩基での結合力を次のように計算可能であると仮定する

配列が一致した場合に、結合力があり、一致しない場合には結合力0とする

G-C対は、A-T対の $\alpha$ 倍の結合力を有する

プローブ内位置での結合力差異は、位置に関する一次式で表現可能

$$pos(i) = \beta \cdot (i - 13) + 1$$

不一致部分から、n-mer連続して一致した箇所から結合力が発生する

$$Bind = \sum_{i=1}^{25} s_i \cdot \alpha_i \cdot pos(i) \cdot \gamma_i$$

洗浄工程においては、25mer一致以外は洗い流されると仮定する。  
検証を迅速に行うために仮定する

Copyright©2008-2009 NTT DATA Corporation

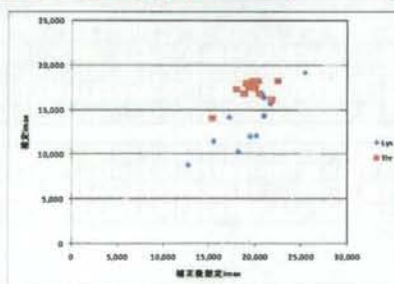
## 2.2.15. 離脱速度比例定数（結合強度）の算出（試算）



不一致と5個以上離れている場合に結合力を発生すると仮定した。  
ただし、ThrとLysの間では、9個以上連続で一致することはなく、Probe末端での結合が反映されている

G-C比率 0.54 GC塩基がAT塩基に対する結合力  
結さ -0.01 5末端と5末端の間の結合力差異  
 $\gamma_{end}$  1末端における結合力  
 $\gamma_{mid}$  0.82 中間における結合力

G0ペアの方が弱いという結果になった。  
GSC第1濃度、第2濃度の特殊な関係の可能性もある



正の相関がみられ、阻害が弱く見積もられているようである。

Copyright©2008-2009 NTT DATA Corporation

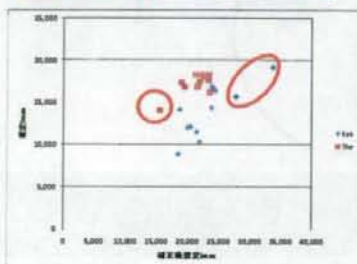
## 2.2.15. 離脱速度比例定数（結合強度）の算出（試算）



不一致と4個以上離れている場合に結合力を発生すると仮定した。  
ただし、ThrとLysの間では、9個以上連続で一致することはなく、Probe末端での結合が反映されている

GC含量 0.431 GC塩基がAT塩基に対する結合力  
 補正 0.042 3'末端と5'末端の間の結合力差  
 7' end 1'末端における結合力  
 7' mid 0.825 7'末端における結合力

GCペアの方が弱いという結果になった。  
GSC第1濃度、第2濃度の特殊な関係の可能性もある



第3プローブを除き補正後の最大値が一定の範囲に収まっている。

Copyright©2008-2009 NTT DATA Corporation

## 2.2.15. 離脱速度比例定数（結合強度）の算出（試算結果検討）



離脱速度の計算をGSCの飽和実験結果を用いて実施した。

GSCの2RNA間である程度の結果が出た。

改善すべき点が見つかった。

GCペアが、ATペアよりも結合力が弱い

対象RNAを増やし、配列条件を増やす必要がある

範囲に収まらない組み合わせが存在した

結合力の考え方（塩基間の強度の合算）

外れることの可能性を考えるためには、熱力学的偶然性で評価すべきと思われる。

Washing時の離脱速度

25mer以外でも残るとすれば、現象を説明しやすいか検討する

Copyright©2008-2009 NTT DATA Corporation